UNCLASSIFIED

AD NUMBER
AD456962
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; 16 OCT 1964. Other requests shall be referred to U.S. Navy Electronic Laboratory, San Diego, CA 92152.**
AUTHORITY
USNEL per ltr, 16 Sep 1965

THIS PAGE IS UNCLASSIFIED

UNCLASSIFIED

AUTOMATIC DISTRIBUTION-FREE STATISTICAL SIGNAL DETECTION

A review and evaluation of available techniques

C. B. Bell

Research Report

16 October 1964

U. S. NAVY ELECTRONICS LABORATORY, SAN DIEGO, CALIFORNIA 92152 • A BUREAU OF SHIPS LABORATORY

NEL REPORT 1245

THE PROBLEM

Develop new radar techniques based upon statistical methods which are applicable to automatic radar systems with increased detection capabilities. The specific phase reported here is a review and evaluation of some of the distribution-free statistical techniques which can possibly be used in signal detection and CCM applications.

RESULTS

The use of distribution-free detectors is quite feasible since:

1. Several of the distribution-free detectors are uniformly better than the parametric detectors for certain classes of Gaussian noise.

2. Several simpler distribution-free detectors are almost as efficient as their parametric counterparts.

RECOMMENDATIONS

1. Initiate procedures to implement the design and construction of one or more of the better distribution-free detectors.

2. Investigate better comparison for the various types of distribution-free detectors.

3. Initiate simulated tactical tests of distribution-free detectors as soon as possible.

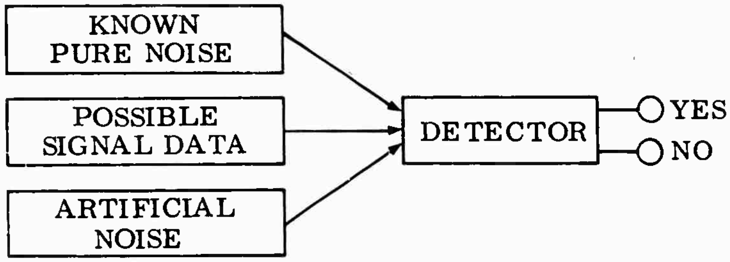
4. Review current research in distribution-free statistics for additional applications to distribution-free detectors.

ADMINISTRATIVE INFORMATION

Work was performed under SF 001 0205, Task 6072 (NEL D1-17-5, formerly D1-7, D1-8, and D1-9) as authorized by Bureau of Ships letter C-9670/2, ser 684B-028, of 12 June 1962. This report covers work from August 1961 to June 1964, and was approved for publication 16 October 1964.

Monte Carlo computations were obtained through the assistance of G. Dillard and R. Worley and the programming of R. Arenz. P. Chase assisted with other calculations.

CONTENTS



I. INTRODUCTION

A. GENERAL

Signal detection, as employed in automatic radar (or sonar) systems, is treated as a problem of testing statistical hypotheses. In its simplest form, this reduces to a choice between signal or no signal. The choice is inferred from the data collected by the radar (or sonar) system through the application of a decision rule or process. When the statistical nature of the data is known for the signal and no-signal cases, then a conventional procedure, such as the Neyman-Pearson test or the sequential probability ratio test, may be used.

When the noise or signal-plus-noise distribution is unknown, the afore-mentioned detection methods present two major difficulties. If the underlying statistics are known to be unknown, then the detection process fails to materialize; if they are incorrectly assumed, the consequent detector performance can be absurd.

Distribution-free methods of signal detection can be employed whenever the underlying signal and noise distributions are unknown. This can occur, for instance, in a jamming or countermeasures environment, or possibly in less hostile circumstances where the lack of physical knowledge, or where *a priori* knowledge of physical considerations indicates uncertainties in data statistics in either a temporal or spatial sense.

The study reported here involved a number of distribution-free tests which offer some promise of improved performance under circumstances such as those previously discussed. Information related to the required data processing, distribution-free detector characteristics, and comparison of distribution-free detectors and classical detectors is presented. The physical or tactical considerations related to the choice of a distribution-free detector for any particular application is not discussed and must be considered at some later time. Indeed, the relation is strengthened and generalized by the avoidance of particular and restrictive applications. Specific radar systems are not considered but, instead, a general approach is used.

The general approach consists of treating the radar target and other factors external to the radar system proper as a "transmission medium" through which the desired signal is not only received but is obscured by noise. The signal might result from the target's reflection of radar energy derived from the electromagnetic waves from the radar transmitter, or the target itself might be the primary energy source. The noise might be additive such as atmospheric noise or artificial interference, or might consist of additive as well as signal-perturbing components. Three models are presented.

Detection is accomplished by means of a device which receives known pure noise, possible signal (PS) data, and/or artificial noise, and from these inputs "decides" (YES or NO) whether or not a signal is present. In making this decision, the detector is capable of producing two types of erroneous outputs: (1) false alarm (FA), if NO signal is present and it decides YES; and (2) false dismissal (FD) if YES a signal is present and it decides NO.

The probability that the detector will produce a false alarm (PFA) will be designated α; β will denote the probability that the detector will produce a false dismissal (PFD).

The perfect detector, which would produce $\alpha = 0$ and $\beta = 0$, has not yet been constructed. However, there is considerable

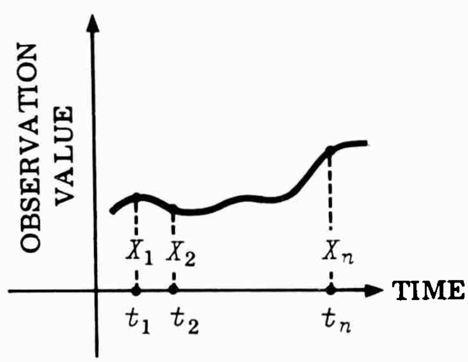
interest in the so-called ideal detector, which minimizes the probability of an error, either FA or FD

$$p\beta + (1-p)\,\alpha$$

where p is the *a priori* probability that a signal is present. To construct the ideal detector it would be necessary to know the probability p.

Interest will be directed here to the distribution-free detector which in a subsequent section will be defined and compared with the uniformly minimum PFD detector which, for fixed PFA α, minimizes PFD β. It is first necessary to consider some of the statistical properties of pure noise and of noise-plus-signal.

B. STATISTICAL PROPERTIES OF NOISE AND SIGNAL



Throughout it is assumed that the PS data received by the detector consist of discrete observations $X_1, X_2, \cdots, X_n$ of a continuous time process at times $t_1, t_2, \cdots, t_n$; and that the X's are measurements of voltage, current, power, etc. Further, the following restrictions will be imposed on the succeeding analysis.

1. $X_1, X_2, \cdots, X_n$ are statistically independent random variables;

2. $X_1, X_2, \cdots, X_n$ have a common strictly increasing continuous cumulative probability function (cpf) F such that $F(x) = P(X_i \leq x)$.

(It is clear from the definitions that a cpf F is monotone non-decreasing so that $F(-\infty) = 0$ and $F(+\infty) = 1$.)

3. If the PS data are pure noise, then $F = F_0$; if the PS data are noise-plus-signal, then $F = F_1 \neq F_0$.

The first two assumptions are equivalent to the statement that $X_1, \cdots, X_n$ constitute a random sample from a population with cpf F. More specifically, assumption 2 states that the observed process is stationary or time-invariant. Under certain circumstances this assumption may not hold for long periods of time, but should be approximately valid for relatively short periods.

Assumption 1, the assumption of independence, which is valid in some situations, also represents an approximation in many real situations. The approximation will be quite good if, for example, the continuous time process is a normal or Gaussian process in which the covariance function decreases to zero rapidly relative to the time intervals ($t_i - t_{i-1}$) between observations.

Assumption 3 is, of course, essential to the whole concept of statistical methods in detection. If the statistical properties of noise and noise-plus-signal are not different, then it will not be possible to design a detector using statistical procedures except with the undesirable property that $\alpha + \beta = 1$, i.e., with (PFA) = 1-(PFD).

As previously mentioned, the detector will have as inputs under certain circumstances not only PS data, but also pure noise as well as artificial noise. The pure-noise observations will be represented as $Y_1, Y_2, \cdots, Y_n$ and will also satisfy assumptions 1 and 2 and, of course, assumption 3 with $F = F_0$. Further, whenever PS data and pure noise are simultaneously available, they will be assumed statistically independent.

In order to facilitate computations or to achieve certain exact PFA values α, it will sometimes be feasible to introduce artificial noise observations $W_1, W_2, \cdots, W_N$. These random variables will be assumed to satisfy assumptions 1 and 2 and to have some continuous cpf H, which is, in general, not equal to F_0 nor to F_1.

With these preliminary statistical considerations it is now possible to introduce the three major statistical models of detectors to be considered in subsequent sections.

C. THREE DISTRIBUTION-FREE DETECTOR MODELS

Before proceeding to a study of distribution-free models one should point out that in the event that both F_0 and F_1 are known, and PS data $X_1, \cdots, X_n$ are available, it is possible from well-known statistical considerations to design a uniformly minimum PFD detector, i.e., one which for fixed PFA α minimizes PFD β. Of course, such a detector would not in general have this minimizing property for other F_0's and F_1's and, hence, would not be practicable for use except when the noise and noise-plus-signal cpf's are approximately F_0 and F_1, respectively.

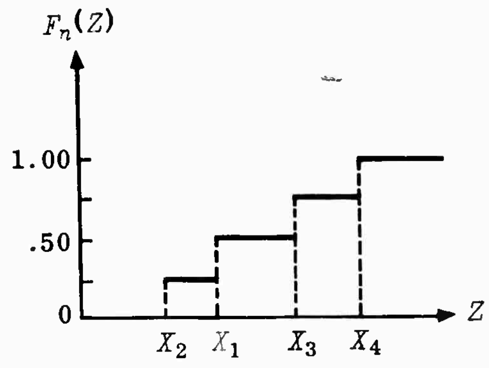
Detectors whose performance (PFA, etc.) does not depend on a specific fixed continuous F_0 will be called distribution-free detectors. The models of distribution-free detectors to be studied here are based on three reasonable types of statistical situations:

MODEL I: The pure-noise cpf F_0 is known and there are available PS data $X_1, X_2, \cdots, X_n$ whose cpf, F_1, is unknown.

MODEL II: Both the cpf's F_0 and F_1 are unknown, but there are available PS data sample $X_1, X_2, \cdots, X_n$ and pure-noise sample $Y_1, \cdots, Y_m$. (This model was primarily developed by J. Capon; see reference 1 in list at end of report.)

MODEL III: F_0, F_1 and a pure noise sample are not available, but the scanned regions can be divided into subregions and each PS data observation can be classified according to its subregion of origin as follows:

1st Subregion:	$X_{11}, X_{12}, \cdots, X_{1,n_1}$
2nd Subregion:	$X_{21}, X_{22}, \cdots, X_{2,n_2}$
. . .	. . .
k^{th} Subregion:	$X_{k1}, X_{k2}, \cdots, X_{k,n_k}$



Sample cpf for four observations.

In making decisions on the basis of Models I, II, and III, it is customary to introduce the well-known statistical concept of sample cpf. For the observations $X_1, X_2, \cdots, X_n$, the sample cpf F_n is a step function such that $F_n(z) = \frac{1}{n}$ (number of X's $\leq z$) for all z, as illustrated here for $n = 4$.

In terms of the sample cpf's the basic approaches of the three models are as follows.

1. A Model I detector compares the PS sample cpf F_n with the pure-noise cpf F_0 and decides YES if and only if F_n is "too far" from F_0. Essentially, each different Model I detector uses a different definition of "too far." These ideas are developed more fully in Section II.

2. A Model II detector compares the sample cpf F_n of the PS data $X_1, \cdots, X_n$ with the sample cpf F_m of the pure-noise observations $Y_1, Y_2, \cdots, Y_m$, and employs extensions of the ideas mentioned above for Model I. In Section III this is developed in more detail.

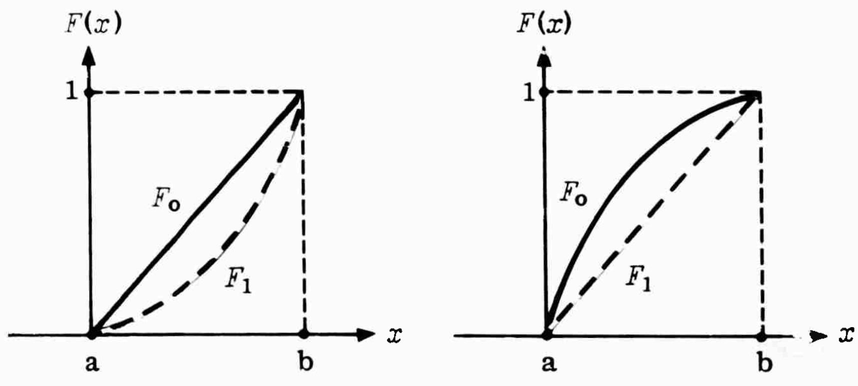
3. A Model III detector compares the sample cpf's F_{n_1}, $F_{n_2}, \cdots, F_{n_k}$ of the observations from the k subregions. This model is treated in more detail in Section IV.

Before discussing the three models in detail it seems worthwhile to evolve a formula for producing distribution-free detectors.

D. CHARACTERIZATION OF A USEFUL CLASS OF DETECTORS

Paralleling known statistical techniques, one can characterize the class of Model I detectors by imposing a natural restriction on the PFD β. Note first that if both the pure-noise cpf F_o and the noise-plus-signal cpf F_1 are normal (or Gaussian) cpf's, then it is known (see Section II D) that for each PFA α, $\beta = \beta(F_o, F_1)$ depends only on the function $F_o(F_1^{-1})$, i.e., the composition of F_o and the inverse, F_1^{-1}, of F_1. Extending this idea to the distribution-free case one makes the following definition.

Definition 1. A detector is called <u>SDF</u> (strongly distribution-free) if for each PFA α, each pure-noise cpf F_o, and each noise-plus-signal cpf F_1, the PFD, $\beta = \beta(F_o, F_1)$, depends only on the function $F_o(F_1^{-1})$ for all strictly monotone continuous cpf's F_o and F_1. This would mean, for example, that an SDF detector would for each fixed PFA α have the same PFD β in both of the cases illustrated. This so since in both of these cases $F_1 = F_o^2$ and $F_o\left(F_1^{-1}(u)\right) = u^{\frac{1}{2}}$ for all u.



Further one needs

Definition 2. A Model I detector is symmetric if its decisions do not depend on the chronological order of the $X_1, \cdots, X_n$. (Of course, such a detector is not sequential.)

From the statistical results of Birnbaum and Rubin[2] and those of Bell[3,4] it is known that

Theorem 1. An SDF symmetric Model I detector must base its YES-NO decisions on a statistic of the form $\psi[F_o(X_1), \cdots, F_o(X_n)]$, where ψ is a symmetric function and F_o is the known continuous pure-noise cpf.

Table 1 gives some examples of SDF symmetric Model I detectors, with $X(1) < X(2) < \cdots < X(n)$ representing the ordered values of $X_1, X_2, \cdots, X_n$.

Table 1. Some SDF Symmetric Model I Detectors

Name of Detector	Decision YES iff (if and only if)
Kolmogorov D_n	$\max_i \left\{ \max \left[F_n\left(X(i)\right) - \frac{i-1}{n}, \frac{i}{n} - F_n\left(X(i)\right) \right] \right\} > d_1(n, \alpha)$
Fisher π	$-2 \Sigma \ln F_o(X_i) < d_2(n, \alpha)$
$\overline{U}$	$\Sigma F_o(X_i) < d_3(n, \alpha)$
Sign Q'_p	$F_n\left(F_o^{-1}(p)\right) > d_4(n, \alpha, p)$
Sherman $\tilde{S}_n$	$\Sigma \left\| F_o\left(X(i)\right) - F_o\left(X(i-1)\right) - \frac{1}{n+1} \right\| > d_5(n, \alpha)$

For Model II detectors it is feasible also to impose the SDF condition, i.e., that $\beta(F_o, F_1)$ depend only on $F_o(F_1^{-1})$. However, in giving the characterization of Model II detectors one must recall the concept of ranks $R(X_1), \cdots, R(X_n)$; $R(Y_1), \cdots, R(Y_m)$ in the combined sample.

Example 1. If the pure-noise observations are $Y_1 = 2.1$, $Y_2 = 3.2$, $Y_3 = 0.9$, and the PS data are $X_1 = 0.8$, $X_2 = 2.5$, $X_3 = 1.6$, $X_4 = 3.1$, then one has

1. as order statistics for the Y-sample $Y(1) = 0.9$, $Y(2) = 2.1$, $Y(3) = 3.2$;

2. as order statistics for the X-sample $X(1) = 0.8$, $X(2) = 1.6$, $X(3) = 2.5$, $X(4) = 3.1$;

3. as the order in the combined sample $XYXYXXY$; and, hence,

4. that the X's have ranks 1, 3, 5, and 6 in the combined sample, while the Y's occupy ranks 2, 4, and 7. More specifically, $R(X_1) = 1$, $R(X_2) = 5$, $R(X_3) = 3$, $R(X_4) = 6$, $R(Y_1) = 4$, $R(Y_2) = 7$, $R(Y_3) = 2$.

Further one needs the definitions:

Definition 3. A Model II detector is a rank detector if its decisions depend only on the ranks of the X's (and/or Y's; and not on their numerical values).

Definition 4. A Model II detector is SWS (samplewise symmetric) if its decisions depend neither on the chronological order of $X_1, \cdots, X_n$ nor on the chronological order of $Y_1, \cdots, Y_m$.

It is known from statistical considerations[3,5] that each rank detector is an SDF, SWS detector. However, one needs the converse of this result, and to that end must introduce an additional restriction. This mathematical condition first considered by Scheffé[5,6] is satisfied by all of the Model II detectors known to the author; but the full significance of the condition is not evident to the author.

Definition 5. A Model II detector is a Scheffé detector if its decisions are based on a statistic $T = T(X_1, \cdots, X_n ; Y_1, \cdots, Y_m)$ such that the boundary of the inverse image $T^{-1}(B)$ of each Borel set B has probability 0 with respect to each continuous (power) measure.

This Scheffé condition is quite complex, and need not concern the reader other than as a condition to guarantee the following theorem.

Theorem 2. An SDF, SWS Model II Scheffé detector must be a rank detector, i.e., must base its YES-NO decisions solely on the ranks of the $X_1, \cdots, X_n$ and $Y_1, \cdots, Y_m$ in the combined sample.

Typical examples of Model II rank detectors are given in table 2.

Table 2.

Name	Decision: YES iff
Cramér-von Mises	$\int \left[F_n(z) - F_m(z)\right]^2 d\left[\frac{nF_n(z) + mF_m(z)}{n+m}\right] > d_1$
Mann-Whitney-Wilcoxon	$\frac{1}{n} \Sigma R(X_i) > d_2$

For a Model III detector one recalls that there are available PS data from each of k subregions;

$$\begin{array}{l} X_{11},\ X_{12}, \cdots, X_{1,n_1} \\ X_{21},\ X_{22}, \cdots, X_{2,n_2} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ X_{k1},\ X_{k2}, \cdots, X_{k,n_k} \end{array}$$

Model III detectors can be characterized (as can Model II) by requiring the SDF, SWS, and Scheffé condition.

Theorem 3. An SDF, SWS Model III Scheffé detector must be a rank detector, i.e., must base its YES-NO decisions solely on the ranks $R(X_{11}), \cdots, R(X_{k,n_k})$ of the samples from the k subregions.

Typical examples of Model III rank detectors are given in table 3.

Table 3. Some Model III Rank Detectors

Name	Decision: YES; iff
Kolmogorov	$\max \sum_{i=1}^{k} n_i \left[F_{n_i}(x) - \bar{F}(x)\right]^2 > d_1$ where $\bar{F} = N^{-1}\Sigma n_i F_{n_i}$ and $N = \Sigma n_i$
Kruskal-Wallis	$\frac{12}{N(N+1)} \Sigma \frac{R_i^2}{n_i} - 3(N+1) > d_2$ where $R_i = \sum_j R(X_{ij})$
Mosteller-Tukey	Length of last run $> d_3$

Before summarizing this section on characterization one should make two relevant comments.

1. Whereas the rank detectors correspond to a large class of 2-sample and k-sample statistics in common usage, one important class of detectors has been omitted. These are the Pitman type, which are not rank detectors as defined here.*

2. In subsequent sections artificial noise will sometimes be introduced into the Model II and Model III detectors in order to achieve previously stated aims. For Model II, the individual PS data observations $X_1, \cdots, X_n$ and the individual

* Their statistics are discussed in ref. 7 and in ref. 8, p. 489.

pure-noise observations $Y_1, \cdots, Y_m$ will be replaced by that one of the artificial noise observations $W_1, W_2, \cdots, W_N$ ($N = n + m$) with the same rank. For example, if X_2 has rank 5 in the combined X-Y sample then it will be replaced by that W_9 iff W_9 has rank 5 in the W sample. Thus, the rank ordering of the transformed observations is exactly that of the original observations. Consequently, a detector which bases its decisions on the W's is also a rank detector. Analogous statements are also valid for Model III detectors. These ideas are treated in more detail in Sections III and IV.

Table 4 summarizes the results of this section.

Table 4. Summary for Model I, II, and III Detectors

	Model I	Model II	Model III
Cpf's known	Pure-noise cpf F_o	None	None
Data available	PS data: $X_1, \cdots, X_n$	PS data: $X_1, \cdots, X_n$ Pure-noise data: $Y_1, \cdots, Y_m$	Regional PS data: $X_{11}, X_{12}, \cdots, X_{1,n_1}$ $X_{21}, X_{22}, \cdots, X_{2,n_2}$ . . . $X_{k1}, X_{k2} \cdots, X_{k,n_k}$
Characterization assumptions	SDF, symmetric	SDF, symmetric Scheffé	SDF, symmetric Scheffé
Structure of detector statistic	$\psi_1\left[F_o(X_1), \cdots, F_o(X_n)\right]$	$\psi_2\left[R(X_1), \cdots, R(Y_m)\right]$	$\psi_3\left[R(X_{11}), \cdots, R\left(X_{k,n_k}\right)\right]$

Since the number of statistics having the structures indicated is infinite for each model above, some method of narrowing the classes of possibilities should be devised. This is treated in the next section.

E. GOODNESS CRITERIA FOR DETECTORS

Since the number of possible detectors is infinite it would seem reasonable that not all of them are equally "good." Consequently, many of these possible detectors can be eliminated from consideration by the imposition of reasonable goodness criteria. In other words, consideration will be restricted here to those detectors which satisfy one or more practical, reasonable goodness criteria.

Since a detector is essentially a decision-maker, the ultimate objective is the design of a detector which minimizes loss in some appropriate sense. The losses in the given models are, of course, of two possible types:

> L(FD), the loss incurred by an FD (false dismissal), i.e., by deciding NO when there is signal present; and
>
> L(FA), the loss incurred by an FA (false alarm), i.e., by deciding YES where there is only pure noise.

If the detection process is to be repeated in a large number of locations or for a large number of times in a given location, then (as is assumed throughout this work) statistical considerations such as expected loss or risk, PFA α, PFD β, etc., are important, and some consideration of

> p, the *a priori* probability of noise-plus-signal, is necessary.

As previously mentioned the most desirable detector would be the

1. perfect detector, for which $\alpha = 0$ and $\beta = 0$.

Of course, such a detector rarely, if ever, exists. However it is possible to achieve $\alpha = 0$ by always deciding NO; or $\beta = 0$ by always deciding YES. These are

2. trivial detectors.

A more practical detector can be obtained for large classes of pure noise and noise-plus-signal cpf's if p, L(FD) and L(FA) are known. Such a detector is a

3. Bayesian detector, which minimizes the risk

$$p\beta \, L(FD) + (1 - p)\, \alpha \, L(FA).$$

In the event that the loss function L is not known or if L(FD) = L(FA), one may settle for an

4. ideal detector, i.e., one which minimizes $p\beta + (1 - p)\alpha$.

In many practical cases, of course, neither p nor the losses L(FA) and L(FD) are known, and one is therefore led to criteria based solely on PFA α and PFD β. From one point of vicw the "best" such detector is the

5. uniformly minimum (UM) PFD detector, which, for fixed α, gives minimum PFD β for each noise-plus-signal cpf F_1.

It is well known from the Neyman-Pearson result that for each fixed pure-noise cpf F_0 and noise-plus-signal cpf F_1, there exists a detector which for fixed α minimizes β. However, in general, no one detector will minimize β for a "large" class of cpf's F_0 and F_1; and a uniformly minimum PFD detector will exist only in special cases. Hence, one needs other reasonable goodness criteria.

A detector D is called an

6. admissible dctector if there is no other detector with uniformly smaller β, i.e., if for each other detector D_2 there is some pure-noisc cpf F_2 such that $\beta(D_1, F_2) \leq \beta(D_2, F_2)$; an

7. unbiased detector if $\alpha + \beta \leq 1$ for all pure-noise cpf's F_0, i.c., if $P(\text{YES} \mid \text{pure noise}) \leq P(\text{YES} \mid \text{noise-plus-signal})$; and a

8. consistent detector if for each fixed PFA α, $\lim \beta = 0$ as the sample size(s) increase(s) without bound.

Criteria 6, 7, and 8 are quite reasonable. If one detector has smaller β than a second detector under all circumstances, then one would prefer the first detector. Hence, admissibility is a reasonable property to require of detectors.

Also, it seems reasonable that one would want to restrict consideration to detectors which, for fixed α, have increasingly small PFD β as the quantity of data increases.

Finally, as is the opinion of many who engage in applications of statistics, it seems desirable to have the detector decide YES more often when there is signal than when there is only pure noise. Such a detector is unbiased.

Up to this point it has been seen that

a. perfect detectors rarely, if ever, exist;

b. trivial detectors can always be constructed;

c. Bayesian and ideal detectors require specialized *a priori* knowledge; and

d. uniformly minimum PFD detectors exist only in special cases.

Further, it can be shown that the requirements of unbiasedness, admissibility, and consistency neither reduce the class of possible detectors to a "reasonable" number nor provide criteria for comparison. Consequently, one is not in a position to construct a "best" detector at this time; nor can one on the basis of the preceding criteria make statements of the form "detector D_1 is 50 per cent better than detector D_2." In an attempt to satisfy this need one introduces the following two concepts:

9. Detector D_1 is max-min-β better than detector D_2 for pure-noise cpf class $\mathcal{N} = \{F_0\}$ and noise-plus-signal cpf class $\mathcal{S} = \{F_1\}$ if, for each F_0 in $\mathcal{N}$,

$$\sup \beta(D_1) \leq \sup \beta(D_2)$$

and

$$\inf \beta(D_1) \leq \inf \beta(D_2)$$

where the suprema and infima are taken over the class $\mathcal{S}$ of noise-plus-signal cpf's.

The rationale for this criterion (essentially conceived by Chapman,[9] and used by Bell, Moser, and Thompson[10]) is as follows. If detector D_1's best performance is better than detector D_2's best performance for specified reasonable classes $\mathcal{N} = \{F_o\}$ and $\mathcal{S} = \{F_1\}$ and if detector D_1's worst performance is better than detector D_2's worst performance for these classes, one says that detector D_1 is "better" than detector D_2 (for these classes). Of course, this rating of detectors depends explicitly on the classes chosen; and, further, one suspects that only in special cases will both $\sup \beta(D_1) \leq \sup \beta(D_2)$ and $\inf \beta(D_1) \leq \inf \beta(D_2)$.

The second of these two concepts is that of asymptotic relative efficiency. Any detector of "feasible" design can be expected to have good performance when there is a strong signal. Consequently, it seems reasonable to compare two detectors on the basis of their relative performances in the presence of increasingly weak signals.

10. The asymptotic relative efficiency (ARE), $A(D_1, D_2)$ of detector D_1 with respect to detector D_2, is defined to be the limit, $\lim \frac{n(D_2)}{n(D_1)}$, of the ratio of their sample sizes as fixed α and β are maintained as the signal strength tends to zero, i.e., as F_1 tends to F_o.

This statistical concept (often attributed to Pitman[7]) has been recently treated by a large number of authors.[11-18] In order for the limit to exist, be unique, and be independent of α and β, a large number of regularity conditions are needed. (These are mentioned in Section II D.) Further, although one can make statements of the form "D_1 is 50 per cent better than D_2" on the basis of ARE, the ARE depends quite heavily on the pure-noise cpf F_o and the

manner in which the signal tends to zero, i.e., the manner in which the noise-plus-signal cpf F_1 tends to F_o.

Consequently, in order to apply criterion 9 and rank a given set of detectors, or to apply 10 and assign a number to each detector, one must specify classes of pure-noise and noise-plus-signal cpf's.

Considerations such as these lead one to consider detectors whose performances are the same for "large" classes of pure-noise and noise-plus-signal cpf's. Such detectors are strongly distribution-free; one repeats the definition of Section I D.

11. A detector is called an <u>SDF detector</u> if for each PFA α, each pure-noise cpf F_o, and each noise-plus-signal cpf F_1, the PFD, $\beta = \beta(F_o, F_1)$, depends only on the function $F_o(F_1^{-1})$ (for all strictly monotone continuous cpf's F_o and F_1).

Very recent statistical results of Bell and Doksum[12] can be adapted to yield two additional goodness criteria for detectors.

Let $\{F_\theta\}$ be a class of noise-plus-signal cpf's such that $\lim_{\theta \to \theta_o} F_\theta = F_{\theta_o} = F_o$, the pure-noise cpf.

12. A detector D_1 is said to be a <u>locally minimum (LM) PFD detector</u> for the class $\{F_\theta\}$ above, if there exists $\epsilon > 0$ such that the PFD $\beta(D_1, F_o) \leq \beta(D_2, F_\theta)$ for all other detectors D_2 and all θ satisfying

$$|\theta - \theta_o| < \epsilon .$$

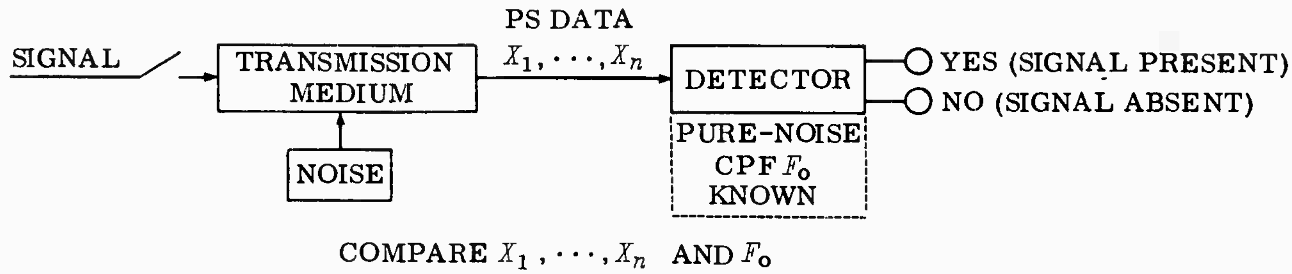
Quite often it is impossible to design such a detector and, for a given class $\{F_\theta\}$ of the form above, one may settle for an

13. <u>almost locally minimum (ALM) PFD detector</u>, which is a detector whose YES-NO decisions are based on a statistic T' satisfying: for every positive δ and η, there exists a positive ϵ such that $P_o(|T' - T(\theta)| > \eta) < \delta$ whenever $|\theta - \theta_o| < \epsilon$, where, for each θ, $T(\theta)$ is the statistic of the <u>uniformly minimum</u> (UM) PFD detector for pure-noise cpf F_o vs. noise-plus-signal cpf F_θ.

It will be seen in Sections II D 1 and III E 1 that for SDF detectors and families of noise-plus-signal cpf's satisfying certain regularity conditions, one can develop explicit formulas for the LM PFD and ALM PFD detectors.

Before beginning a detailed study of the detector models one should mention that there are certain practical considerations, e. g., those of equipment design, and those of calculation and tabulation of the appropriate statistical distributions, which lead one to consider detectors which violate some of the goodness criteria, i. e., detectors which are less "optimal" and more "tractable."

One is now in a position to study the detector models in more detail.



II. MODEL I DETECTORS

As one reealls from Section I D, a Model I detector is one for which there are available both the known pure-noise cpf F_o, and PS data $X_1, X_2, \cdots, X_n$. The decision-making process then consists of comparing the PS data and F_o and deciding YES iff they are in some sense "too far" apart. Further, with the imposition of the very reasonable (Section I D) SDF condition one finds that the SDF Model I deteetor must base its YES-NO decisions on statistics of the form $\psi[F_o(X_1), \cdots, F_o(X_n)]$.

Following the "natural" divisions of SDF statistics one considers three classes of Model I detectors:

1. Sample cpf (SDF Model I) detectors, which employ statistics based on differences of the PS data sample cpf F_n and the pure-noise cpf F_o; one recalls that $F_n(y) = \frac{1}{n}$ (number of X's $\leq y$), the sample proportion of X's which are less than or equal to y.

2. Run-block (SDF Model I) detectors, which deal with the number of PS data valucs $X_1, \cdots, X_n$ which fall in certain preassigned intervals and/or the relative spacing of these values; and

3. Rank-sum (SDF Model I) detectors, whose decisions are based on sums of percentile ranks, $F_o(X_i)$, of the $X_1, \cdots, X_n$.

In the next three sections, general formulas and specific examples of the common types of Model I detectors are given. In later sections the various detectors are compared in terms of the previously mentioned goodness criteria — PFA α, PFD β, ARE, etc.

A. MODEL I SAMPLE Cpf DETECTORS

The detectors discussed in this section are primarily based on generalizations and specializations of three well-known distribution-free statistics whose formulas are given below.

1. Kolmogorov-Smirnov Detector

$$D(n, r, \Psi) = \sup_{x} \left[|F_n(x) - F_0(x)| \sqrt{\Psi(F_0(x))} \right]^r \tag{1}$$

The usual specializations and variations of this statistic are as follows.

$$D_n = \sup_{x} \left[|F_n(x) - F_0(x)| \right]$$
$$= \max_{i} \left\{ \max \left[F_0(X(i)) - \frac{i-1}{n}, \frac{i}{n} - F_0(X(i)) \right] \right\} \tag{2}$$

where $X(1) \leq X(2) \leq \cdots \leq X(n)$ are the ordered values of the PS data $X_1, X_2, \cdots, X_n$;

The two one-sided versions:

$$D^+_n = \sup \left[F_n(x) - F_0(x) \right] = \max_{i} \left[\frac{i}{n} - F_0(X(i)) \right] \tag{3}$$

$$D^-_n = \sup \left[F_0(x) - F_n(x) \right] = \max_{i} \left[F_0(X(i)) - \frac{i-1}{n} \right]; \tag{4}$$

and

$$D'_n = \sup_x \left(\left| F_n(x) - F_o(x) \right| \left\{ [F_o(x)]\ [1 - F_o(x)] \right\}^{-\frac{1}{2}} \right)$$
$$= \max_i \left(\left| \frac{i-1}{n} - F_o(X(i)) \right| \left\{ F_o(X(i))\ [1 - F_o(X(i))] \right\}^{-\frac{1}{2}} \right) \quad (5)$$

which is designed to be sensitive to deviations in the tails of the pure-noise cpf F_o.

2. Cramér-von Mises Detector

$$W(n, r, \Psi) = \int \left\{ \left| F_n(x) - F_o(x) \right| \sqrt{\Psi[F_o(x)]} \right\}^r dF_o(x) \quad (6)$$

The common versions of this statistic are:

$$n\omega^2_n = n \int_{-\infty}^{\infty} [F_n(x) - F_o(x)]^2 \, dF_o(x)$$
$$= \frac{1}{12n} + \sum_{i=1}^{n} \left| \frac{2i-1}{2n} - F_o(X(i)) \right|^2 \quad (7)$$

The two one-sided versions

$$W^+_n = n \int_{-\infty}^{\infty} [F_n(x) - F_o(x)] \, dF_o(x) \quad (8)$$

and

$$W^-_n = n \int_{-\infty}^{\infty} [F_o(x) - F_n(x)] \, dF_o(x) \quad (9)$$

are related by $W^+_n = - W^-_n = n \left[\frac{1}{2} - \frac{1}{n} \sum F_o(X_i) \right]$ and, hence, are both equivalent to the rank-sum statistic $\overline{U} = \frac{1}{n} \sum F_o(X_i)$, which is treated in the following section.

$$W^2_n = n \int_{-\infty}^{\infty} [F_n(x) - F_o(x)]^2 \left[F_o(x)\ (1 - F_o(x))\right]^{-1} dF_o(x)$$

$$= -n - \frac{1}{n} \sum_{i=1}^{n} \Big\{(2i-1) \ln\ F_o(X(i)) + [2(n-i)$$

$$+ 1] \ln\left[1 - F_o(X(i))\right]\Big\} \qquad (10)$$

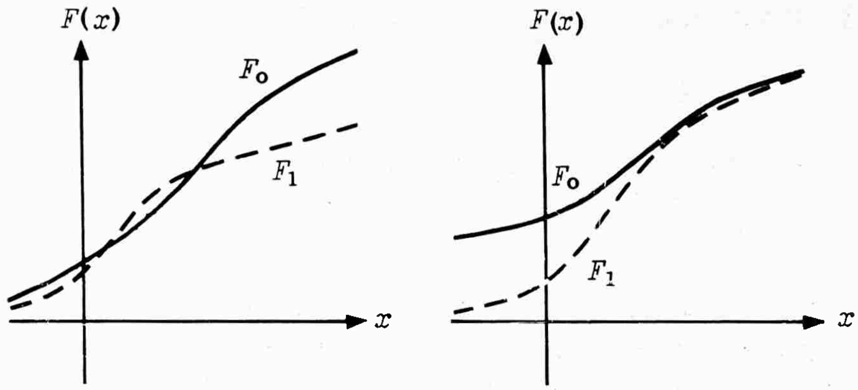
which corresponds to D'_n above and is designed to be sensitive to deviations in the tails of the pure-noise cpf F_o.

On examining the preceding formulas one sees that these statistics are simply generalized distances between the PS data sample cpf F_n and the pure-noise cpf F_o. The Kolmogorov-Smirnov class of detectors is concerned with the maximum of certain functions of the differences $F_n - F_o$, while the Cramér-von Mises class of detectors base their decisions on the integrals of certain functions of $F_n - F_o$. The one-sided versions of both classes of statistics are natural to consider if the noise-plus-signal cpf $F_1 < F_o$ or if $F_1 > F_o$.

One notes that the function $\{F_o(x)\ [1 - F_o(x)]\}^{-\frac{1}{2}}$ diverges as x tends to $\pm\infty$, i.e., as $F_o(x)$ tends to 0 or 1. For this reason

$$|F_n(x) - F_o(x)|\ \ \{F_o(x)\ [1 - F_o(x)]\}^{-\frac{1}{2}}$$

will tend to be large if the absolute difference $|F_n(x) - F_o(x)|$ is large for very small x or for very large x, i.e., differences in the tails of $F_o(x)$ are magnified. Hence, detectors based on equations 5 or 10 are sensitive to noise-plus-signal cpf's F_1 which differ from the pure-noise cpf F_o in the tails.



3. Sign-Quantile Detector

Before giving numerical examples of the use of the detectors just described, one should consider the simplest sample-cpf detector, which is based on the sign statistic and is historically, perhaps, the first distribution-free statistic.

In its rudimentary form a sign detector (sometimes called a threshold detector) is one which decides YES if the number Q' of PS data observations $X_1, \cdots, X_n$ which fall below some preassigned threshold $\xi_p(F_0) = F_0^{-1}(p)$ is "too large." That is,

$$Q' = \sum_{i=1}^{n} \epsilon\,[\xi_p(F_0) - X_i] = \sum_{i=1}^{n} \epsilon\,[p - F_0(X_i)]$$

where ϵ is the degenerate cpf such that

$$\epsilon(z) = \begin{cases} 0 \text{ if } z < 0 \\ 1 \text{ if } z \geq 0 \end{cases}$$

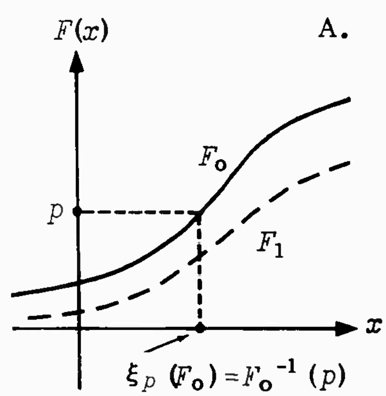
A cursory examination of the situation indicates that in the pure-noise situation approximately $100p$ per cent of the $X_1, \cdots, X_n$ will be equal to or below the threshold $\xi_p(F_0)$; and, of course, approximately $100(1-p)$ per cent of these observations will exceed

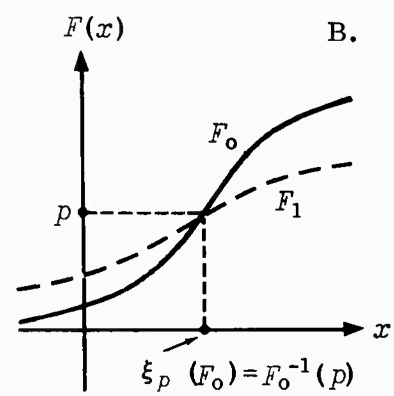
$\xi_p(F_o)$. Consequently, Q' will be too large if

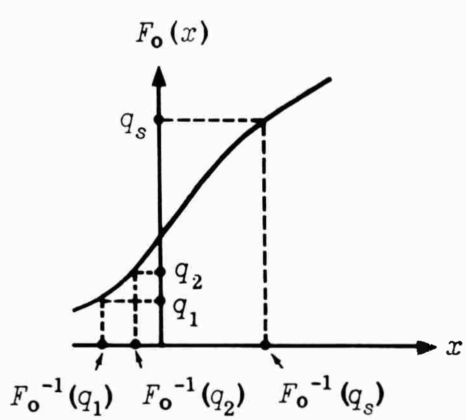
$$F_o(\xi_p(F_o)) - F_n(\xi_p(F_o)) = p - F_n\left(F_o^{-1}(p)\right)$$

is too small, since $nF_n\left(F_o^{-1}(p)\right)$ is the number of $X_1, \cdots, X_n$ which are less than or equal to $\xi_p(F_o) = F_o^{-1}(p)$, the p^{th} quantile of F_o (see illustration, A).

The sign detector, although it will be seen to have certain good properties, does not provide a desirable situation for distinguishing between pure noise and noise-plus-signal if $F_1^{-1}(p) = F_o^{-1}(p)$. In both the pure noise and noise-plus-signal cases, the X's will then fall below the threshold with the same probability, and elementary calculations with show that $\alpha + \beta = 1$, i.e. PFA + PFD = 1 (see illustration, B).







This leads one to consider detectors whose decisions are based not on a single threshold $\xi_p(F_o) = F_o^{-1}(p)$, but on "thresholds" $\xi_{q_1}(F_o)$, $\xi_{q_2}(F_o), \cdots$, $\xi_{q_s}(F_o)$, as diagrammed.

Let $0 = q_0 < q_1 < q_2 < \cdots < q_{s-1} < q_s = 1$ be an arbitrary finite sequence of probabilities. Then

$$\left(F_0^{-1}(q_0),\ F_0^{-1}(q_1)\right],\ \left(F_0^{-1}(q_1),\ F_0^{-1}(q_2)\right],\cdots\cdots,\ \left(F_0^{-1}(q_{s-1}),\ F_0^{-1}(q_s)\right)$$

is a partition of the real numbers into s nonoverlapping intervals on the real line. If there is only pure noise, i.e., if $F = F_0$, then for the i^{th} interval $J_i = \left(F_0^{-1}(q_{i-1}),\ F_0^{-1}(q_i)\right]$,

1. its probability is $P_{F_0}(J_i) = q_i - q_{i-1} = p_i$
2. the expected number of PS data points $X_1, \cdots, X_n$ to fall in the interval is $n(q_i - q_{i-1}) = np_i$
3. the actual number of PS data points in the ith interval $n_i = n\left[F_n\left(F_0^{-1}(q_i)\right) - F_n\left(F_0^{-1}(q_{i-1})\right)\right]$; and
4. each of quantities $\left|F_n\left(F_0^{-1}(q_i)\right) - F_n\left(F_0^{-1}(q_{i-1})\right) - (q_i - q_{i-1})\right| = \left|\frac{n_i}{n} - p_i\right|$ is in some sense a measure of the deviation of the PS sample $X_1, \cdots, X_n$ from F_0.

For these reasons one is led to the following generalization of the original sign statistic.

For the <u>sign-quantile</u> detector,

$$Q(n;\ q_1, \cdots, q_{s-1};\ r;\ \Psi)$$

$$= \sum_{i=1}^{s} \left\{\left|\frac{n_i}{n} - p_i\right| \Psi\left[\frac{n_i}{n}, p_i\right]\right\}^r \tag{11}$$

where for $i = 1, 2, \cdots, s$,

$$n_i = n\left[F_n\left(F_0^{-1}(q_i)\right) - F_n\left(F_0^{-1}(q_{i-1})\right)\right] \text{ and } p_i = q_i - q_{i-1}.$$

On setting $s = 2$, $q_1 = p$, $r = 1$ and $\Psi = 1$, one obtains

a. TWO-SIDED SIGN DETECTOR

$$Q(n;\, p;\, 1;\, 1) = 2\left|\frac{n_1}{n} - p\right| = 2\left|\frac{1}{n}\sum_{i=1}^{n} \epsilon[\,p - F_0(X_i)\,] - p\right|$$

$$= \frac{2}{n}\,\left|Q' - np\right| \tag{12}$$

where Q' and ϵ are, respectively, the sign statistic and degenerate cpf defined above and given below.

b. SIGN DETECTOR

$$Q'_p = \sum_{i=1}^{n} \epsilon[\,p - F_0(X_i)] = \sum_{i=1}^{n} \epsilon[\xi_p(F_0) - X_i\,] = n_1 \tag{13}$$

(Note: When $p = 0.50$, $Q'_{.50}$ is sometimes referred to as the median statistic.)

The most common statistic for $s \geq 2$, i.e., for more than one "threshold," is the following, which is sometimes called a "chi-square" statistic because of its limiting distribution.

c. EXTENDED (MEDIAN) DETECTOR

$$\tilde{Q} = \sum_{i=1}^{s} \frac{(n_i - np_i)^2}{np_i}\text{, which is the specialization of} \tag{14}$$

$Q(n;\, q_1, \cdots, q_{s-1}, r;\, \Psi)$ for which $r = 2$ and $\Psi(u, v) = \sqrt{n/v}$.

In the event that $p_i = 1/s$ or, equivalently, $q_i = i/s$, then one has

$$\widetilde{Q}_s = \frac{1}{ns}\sum_{i=1}^{s}(sn_i - n)^2 = \frac{s}{n}\sum_{i=1}^{s}\left(n_i - \frac{n}{s}\right)^2 \tag{15}$$

From one point of view the extended median $\widetilde{Q}$ gives undue weight to intervals $\left(F_0^{-1}(q_{i-1}),\ F_0^{-1}(q_i)\right] = \left(\xi_{q_{i-1}}(F_0),\ \xi_{q_i}(F_0)\right]$ for which $p_i = q_i - q_{i-1}$ is small because of the denominator np_i. In an attempt to adjust this situation one considers a statistic Q with $r = 2$, and $\psi(u, v) = (\sqrt{u} + \sqrt{v})^{-1}$

d. MATUSITA DETECTOR

$$\hat{Q} = 4n\sum_{i=1}^{s}\left(\sqrt{n_i/n} + \sqrt{p_i}\right)^2 = 4n\sum_{i=1}^{s}\left[\left|\sqrt{n_i/n} - p_i\right| \cdot \left(\sqrt{n_i/n} + \sqrt{p_i}\right)^{-1}\right]^2 = 8n - 8\sqrt{n}\sum_{i=1}^{s}\sqrt{n_i p_i} \tag{16}$$

which reduces in the case $p_i = 1/s$ to

$$\hat{Q}_s = 8n\left[1 - (ns)^{-\frac{1}{2}}\sum_{i=1}^{s}\sqrt{n_i}\right] \tag{17}$$

Of course, the twelve or so types of Model I sample cpf detectors above do not in any sense exhaust all of the possibilities. However, they do give a more or less comprehensive picture of the sample cpf statistics in current use and, hence, of Model I sample cpf detectors which should be considered.

Examples illustrating the computations involved in the use of several of these statistics are given below. Since the normal distribution and the Rayleigh distribution are, of course, two of the more common noise distributions studied in signal detection, the use of the Kolmogorov-Smirnov and Cramér-von Mises detectors

will be illustrated for a pure-noise cpf F_0 which is $N(0,1)$, i.e., normal with $\mu = 0$ and $\sigma^2 = 1$; and the sign-quantile tests will be illustrated for a pure-noise cpf F_0 which is standard Rayleigh, i.e., $F_0(x) = 1 - \exp(-x^2)$ for $x \geq 0$.

Example 2. Consider the following Model I detection situation.

Pure-noise cpf: Standard Rayleigh, $F_0(x) = 1 - \exp(-x^2)(x \geq 0)$

PFA $\alpha = .05$

PS data: 1.3, .3, .2, 1.9, .8 $\quad n = 5$

D_n detector: Decide YES iff

$$\max_i \left\{ \max \left[F(X(i)) - \frac{i-1}{n}, \frac{i}{n} - F_0(X(i)) \right] \right\} > .5633$$

$n\omega^2{}_n$ detector: Decide YES iff

$$\frac{1}{12n} + \sum_{i=1}^{n} \left[\frac{2i-1}{2n} - F_0(X(i)) \right]^2 > .46136$$

(Note: ".5633" is the 95th percentile of the D_n distribution for $n = 5$ obtained from ref. 19, p. 431. ".46136" is the 95th percentile of the $n\omega^2{}_n$ distribution obtained from A. W. Marshall's table, ref. 20.)

Computations

	Basic Computation			Kolmogorov -Smirnov Computation		Cramér-von Mises Computations	
i	$i/N = F_n(X(i))$	$X(i)$	$F_o(X(i)) =$ 1-exp$[X^2(i)]$	$F_o(X(i)) - \frac{i-1}{N}$	$\frac{i}{N} - F_o(X(i))$	$\frac{2i-1}{2n}$	$\left[F_o(X(i)) - \frac{2i-1}{2n}\right]^2$
1	.2	.2	.04	.04	.16	.10	$(-.06)^2 = .0036$
2	.4	.3	.09	-.11	.31	.30	$(-.21)^2 = .0441$
3	.6	.8	.47	.07	.13	.50	$(-.03)^2 = .0009$
4	.8	1.3	.82	.16	.04	.70	$(.12)^2 = .0144$
5	1.0	1.9	.97	.17	.03	.90	$(.07)^2 = .0049$

For the D_n detector

$$D_n = \max(.04, .16, -.11, .31, .07, .13, .16, .04, .17, .03) = .31.$$

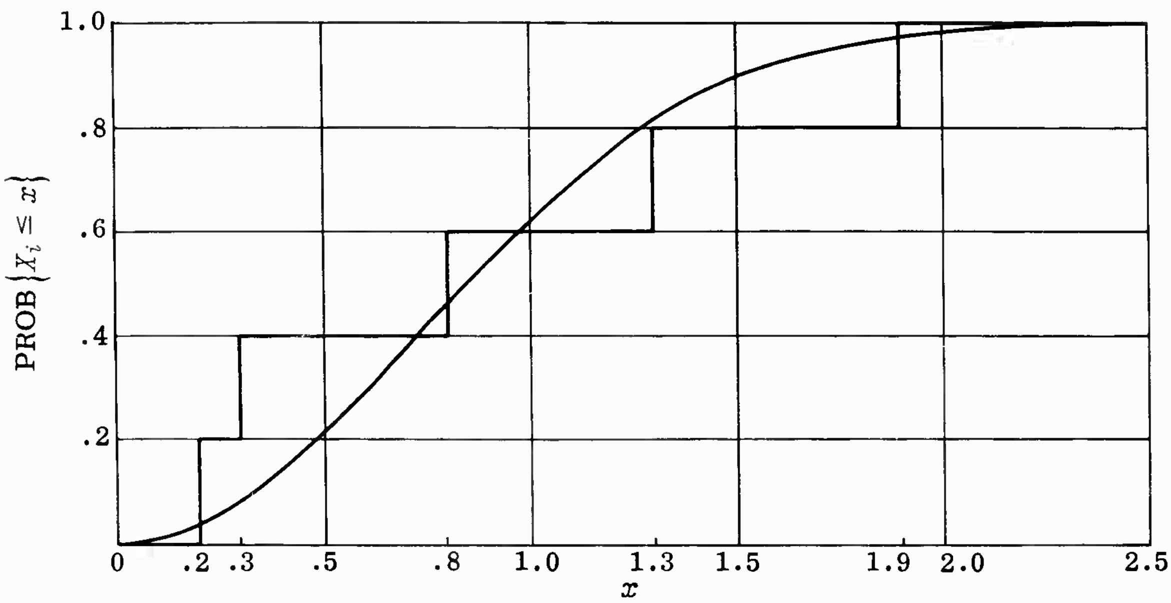
Since $.31 < .5633$, the D_n detector decides NO.

For the $n\omega^2{}_n$ detector

$$n\omega^2{}_n = [(12)(5)]^{-1} + (.0036 + .0441 + .0009 + .0144 + .0049)$$

$= .0167 + .0679 = .0846 < .46136$. Hence, the $n\omega^2{}_n$ detector decides NO.

The approximate value of D_n can be obtained from the graphs of F_n and F_o in the following figure, since D_n is the maximum of the absolute value of the differences $F_n(x) - F_o(x)$. From the graphs one can see that the maximum difference is attained for $x = .3$, and is $F_n(.3) - F_o(.3) = .4 - \left\{1 - \exp\left[-(.3)^2\right]\right\} = .31$.



Rayleigh cpf F_o and sample cpf F_n (see example 2, p. 28).

Note also that from the table above the values of several of the other sample-cpf detector statistics can be computed, as follows:

From the sixth column,

$$D^+_n = \max_i \left[\frac{i}{N} - F_o(X(i))\right] = \max(.16, .31, .13, .04, .03) = .31$$

From the fifth column,

$$D^-_n = \max_i \left|F_o(X(i)) - \frac{i-1}{n}\right| = \max(.04, -.11, .07, .16, .17) = .17$$

From the fourth column,

$$W^+_n = \frac{n}{2} - \sum_{i=1}^{n} F(X(i)) = \frac{5}{2} - (.04 + .09 + .47 + .82 + .97)$$

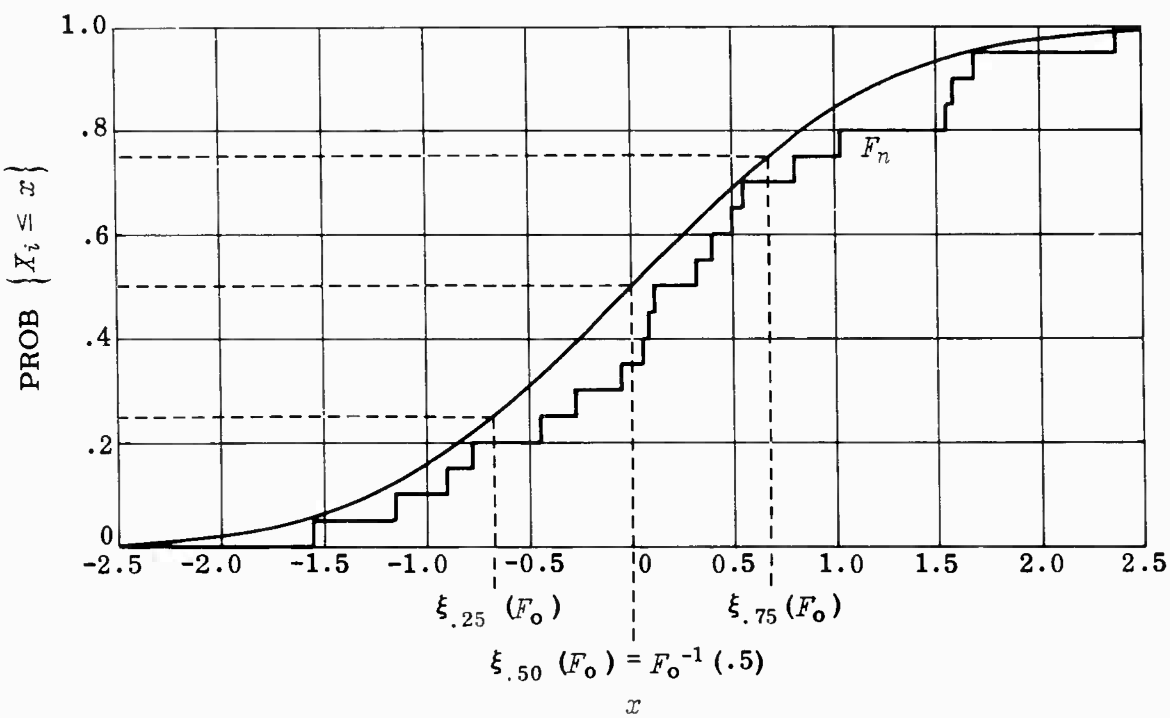
$$= 2.50 - 2.39 = .11, \text{ and}$$

$$W^-_n = - W^+_n = -.11.$$

For an illustration of the use of the sign-quantile detectors one can consider the following example.

Example 3.

Pure-noise cpf: F_0 is $N(0,1)$ (See figure below, in which F_0 and F_n are plotted.)



$N(0,1)$ cpf F_0 and sample cpf F_n (see example 3, p. 31).

PFA $\alpha = .10$ $\qquad$ $n = 20$

(a) $Q'_{.50}$ detector: Decide YES iff $\sum_{i=1}^{20} \epsilon(0 - X_i) \le 6$

(b) $Q'_{.75}$ detector: Decide YES iff $\sum_{i=1}^{20} \epsilon(.675 - X_i) \le 12$

(c) $\tilde{Q}_4$ detector: Decide YES iff $[(20)(4)]^{-1} \sum_{i=1}^{4} (4n_i - 20) \ge 6.25$

(d) $\hat{Q}_4$ detector: Decide YES iff $8(20) \left\{ 1 - [(20)(4)]^{-\frac{1}{2}} \sum_{i=1}^{4} \sqrt{n_i} \right\} \ge 6.25$

PS data: .06, .40, .09, 1.58, -.27, -1.15, 1.68, -.05, -1.55, 1.54, .55, 2.38, .81, -.46, .32, 1.02, -.91, -.78, .50, .11.

Computations: $\xi_{.25}(F_o) = -.675$; $\xi_{.50}(F_o) = 0.00$; $\xi_{.75}(F_o) = .675$

i	$X(i)$	Computations for $Q'_{.50}$		Computations for $Q'_{.75}$		Computations for $\tilde{Q}_4$, $\hat{Q}_4$	
1	-1.55	-		-		1	
2	-1.15	-		-		1	$n_1 = 4$
3	-.91	-		-		1	
4	-.78	-	$7 = Q'_{.50}$	-		1	
							$\xi_{.25}(F_o) = -.675$
5	-.46	-		-		2	
6	-.27	-		-		2	$n_2 = 3$
7	-.05	-		-	$14 = Q'_{.75}$	2	
			$\xi_{.50}(F_o) = 0.00$				$\xi_{.50}(F_o) = 0.00$
8	.06	+		-		3	
9	.09	+		-		3	
10	.11	+		-		3	
11	.32	+		-		3	$n_3 = 7$
12	.40	+		-		3	
13	.50	+		-		3	
14	.55	+	13	-		3	
					$\xi_{.75}(F_o) = .675$		$\xi_{.75}(F_o) = .675$
15	.81	+		+		4	
16	1.02	+		+		4	
17	1.54	+		+		4	
18	1.58	+		+	6	4	$n_4 = 6$
19	1.68	+		+		4	
20	2.38	+		+		4	

(a) For the $Q'_{.50}$ detector: $Q'_{.50} = 7 > 6$ and the $Q'_{.50}$ detector decides NO.

(b) For the $Q'_{.75}$ detector: $Q'_{.75} = 14 > 12$ and the $Q'_{.75}$ decides NO.

(c) For the $\widetilde{Q}_4$ detector: $\widetilde{Q}_4 = (80)^{-1}\left[(16-20)^2 + (12-20)^2 + (28-20)^2 + (24-20)^2\right] = 2.00 < 6.25$ and the $\widetilde{Q}_4$ detector decides NO.

(d) For the $\hat{Q}_4$ detector: $\hat{Q}_4 = (160)\left[1 - (80)^{-\frac{1}{2}}(\sqrt{4} + \sqrt{3} + \sqrt{7} + \sqrt{6})\right] = 4.00 < 6.25$ and the $\hat{Q}_4$ detector decides NO.

In closing this section one should mention two pertinent facts:

1. In none of the cases in Example 3 (p. 31) was the exact PFA α attained, since the exact distributions of the statistics involved are given by:

Theorem 4. If $F = F_0$, (i) Q'_p has binomial distribution with parameters n and p; and (ii) as n tends to infinity, $\widetilde{Q}_s$ and $\hat{Q}_s$ both have asymptotically a chi-square distribution with $(s-1)$ degrees of freedom.

The "6" and "12" for $Q'_{.50}$ and $Q'_{.75}$, provide α's of .0577 and .1018, respectively; $\alpha = .10$ is attainable with these sign statistics only through randomization. Further, for $n = 20$, the "6.25" is only an approximation to the 95th percentile of $\widetilde{Q}_4$ and $\hat{Q}_4$ distributions. A summary of the distribution situation is presented at the end of this section.

2. Up to this point nothing has been developed which would allow one to objectively express a preference of detectors among those already considered. This aspect of the detection problem will be discussed in Section D.

B. MODEL I RUN-BLOCK DETECTORS

The detectors of this type are based on the spacings of the PS data relative to the pure-noise cpf F_0.

It is well-known that

Theorem 5. If $0 = q_0 < q_1 < \cdots < q_{s-1} < q_s = 1$, and there are n PS data observations,

(i) the expected number of observations falling in the interval or cell $\left(\xi_{q_{i-1}}(F_0),\ \xi_{q_i}(F_0)\right] = \left(F_0^{-1}(q_{i-1}), F_0^{-1}(q_i)\right]$ is $n(q_i - q_{i-1}) = np_i$ and

(ii) for integers $k \leq r$, the expected value of the difference of "heights" $F_0(X(r)) - F_0(X(k))$ is $E\left[F_0(X(r)) - F_0(X(k))\right] = \dfrac{r-k}{n+1}$

Hence, when the $(n+1)q_i$ are integers, then each of the quantities $\left|F_0\left(X[(n+1)\,q_i]\right) - F_0\left(X[(n+1)\,q_{i-1}]\right) - p_i\right|$ is in some sense a measure of the deviation of the PS data from F_0. This is also true for the number $\widetilde{E}$ of empty cells among the $\left(F_0^{-1}(q_{i-1}),\ F_0^{-1}(q_i)\right]$. One is therefore led to detectors based on the following statistics.

1. Spacing Detector

$$S(n, q_1, \cdots, q_{s-1}; r, \Psi) = \sum_{i=1}^{s} \left[\left|F_0\left(X[(n+1)q_i]\right) - F_0\left(X[(n+1)q_{i-1}]\right) - p_i\right| \Psi(p_i)\right]^r \tag{18}$$

where each $(n+1)q_i$ is an integer; $0 = q_0 < q_1 < \cdots < q_{s-1} < q_s = 1$; $p_i = q_i - q_{i-1}$, $X(0) = -\infty$ and $X(n+1) = \infty$

The most common versions of this detector are based on the following statistics.

a. *SHERMAN DETECTOR*

$$\tilde{S}_n = S\left(n; \frac{1}{n+1}, \cdots, \frac{n}{n+1}; 1; \frac{1}{2}\right) = \frac{1}{2} \sum_{i=1}^{n+1} \left| F_o(X(i)) - F_o(X(i-1)) - \frac{1}{n+1} \right| \text{; and} \qquad (19)$$

b. *KIMBALL-MORAN DETECTOR*

$$S'_n = S\left(n; \frac{1}{n+1}, \cdots, \frac{n}{n+1}; 2; 1\right) = \sum_{i=1}^{n+1} \left[F_o(X(i)) - F_o(X(i-1)) - \frac{1}{n+1} \right]^2 = \sum_{i=1}^{n+1} \left[F_o(X(i)) - F_o(X(i-1)) \right]^2 - \frac{1}{n+1} \qquad (20)$$

The following example illustrates the computations involved.

Example 4. Pure-noise cpf: $F_o(x) = 1 - \exp(-x^2/4)$ $(x \geq 0)$, PFA $\alpha = .01$, $n = 5$

PS data: 1.6, 3.8, .4, 2.6, .6

<u>$\tilde{S}_n$ detector:</u> Decide YES iff $\tilde{S}_n > 0.57442$*

<u>S'_n detector:</u> Decide YES iff $S'_n > s(5, .01)$

* "0.57442" is the 99th percentile of the $\tilde{S}_n$ distribution for $n = 5$, from ref. 21, p. 448. As of this writing there are no published tables of the distribution of S'_n, although such tables can be constructed with the aid of high-speed computers. Hence, $s(5, .01)$ is not known.

Computations

i	$X(i)$	$F_0(X(i))$	For $\widetilde{S}_n$ $\left\|F_0(X(i)) - F_0(X(i-1)) - \frac{1}{n+1}\right\| = D_i$	For S'_n $[F_0(X(i)) - F_0(X(i-1))]^2$
0	$-\infty$	0		
1	0.4	.04	$\|.04 - 0 - .17\| = .13$	.0016
2	0.6	.09	$\|.09 - .04 - .17\| = .12$	.0025
3	1.6	.47	$\|.47 - .09 - .17\| = .21$	.1444
4	2.6	.82	$\|.82 - .47 - .17\| = .18$	.1225
5	3.8	.97	$\|.97 - .82 - .17\| = .02$	.0225
6	∞	1.00	$\|1.00 - .97 - .17\| = .14$	.0009
			.80	.2944

Therefore $\widetilde{S}_n = .80$ and $S'_n = .29 - .17 = .12$

The $\widetilde{S}_n$ detector then decides YES, since .80 > 0.57442.

The decision of the S'_n detector depends on the unknown $s(5, .01)$.

2. Empty-Cell Detector

The second type of Model I run-block detector is the empty-cell detector, which, as its name suggests, is based on the number of the cells $\left(F_0^{-1}(q_{i-1}), F_0^{-1}(q_i)\right]$ which are empty. (For this detector it is convenient to let $s + 1$ denote the number of cells.)

Before giving the general formula for the statistic involved, one should note that

1. $n_i = n\left[F_n\left(F_0^{-1}(q_i)\right) - F_n\left(F_0^{-1}(q_{i-1})\right)\right]$ is the number of PS observations in the i^{th} cell $J_i = \left(F_0^{-1}(q_{i-1}), F_0^{-1}(q_i)\right]$.

2. $\epsilon\left[F_n\left(F_0^{-1}(q_i)\right) - F_n\left(F_0^{-1}(q_{i-1})\right) - \frac{1}{2n}\right] = \begin{cases} 0 \text{ if } J_i \text{ is empty} \\ 1 \text{ if } J_i \text{ is occupied,} \end{cases}$

where $\epsilon(z) = \begin{cases} 1 \text{ if } z \geq 0 \\ 0 \text{ if } z < 0 \end{cases}$, the degenerate cpf.

3. $\sum_{i=1}^{s+1} \epsilon\left[F_n\left(F_0^{-1}(q_i)\right) - F_n\left(F_0^{-1}(q_{i-1})\right) - \frac{1}{2n}\right]$ is the number of occupied cells; and since s division points $F_0^{-1}(q_1) < \cdots < F_0^{-1}(q_s)$ give $(s+1)$ cells J_i

4. $(s+1) - \sum_{i=1}^{s+1} \epsilon\left[F_n\left(F_0^{-1}(q_i)\right) - F_n\left(F_0^{-1}(q_{i-1})\right) - \frac{1}{2n}\right]$ is the number of empty cells.

Consequently, one defines for the empty-cell detector

$$\tilde{E}(n; q_1, \cdots, q_s) = (s+1) - \sum_{i=1}^{s+1} \epsilon\left[F_n\left(F_0^{-1}(q_i)\right) - F_n\left(F_0^{-1}(q_{i-1})\right) - \frac{1}{2n}\right] \tag{21}$$

One usually chooses the cells to be equiprobable, i.e., $q_i = \frac{i}{s+1}$ and $p_i = \frac{1}{s+1}$; and works either with the statistic

$$\tilde{E}_{s+1} = (s+1) - \sum_{i=1}^{s+1} \epsilon\left[F_n\left(F_0^{-1}\left(\frac{i}{s+1}\right)\right) - F_n\left(F_0^{-1}\left(\frac{i-1}{s+1}\right)\right) - \frac{1}{2n}\right] \tag{22}$$

or with the equivalent statistic

$$\hat{E}_{s+1} = \sum_{i=1}^{s+1} \epsilon\left[F_n\left(F_0^{-1}\left(\frac{i}{s+1}\right)\right) - F_n\left(F_0^{-1}\left(\frac{i-1}{s+1}\right)\right) - \frac{1}{2n}\right] \tag{23}$$

which is the number of occupied equiprobable cells.

Since the distribution of $\hat{E}$ is the one usually tabulated (see ref. 21, p. 454 ff), the detection example below will use $\hat{E}$.

Example 5.

Pure-noise cpf: F_0 is standard Rayleigh, $F_0(x) = 1 - \exp(-x^2)$ $(x \geq 0)$.

PFA $\alpha = .01$

PS data: 1.66, 2.06, 3.16, 1.69, 1.33, .45, 3.26, 1.55, .05, 3.14, 2.15, 3.98, 2.41, 1.14, 1.92, 2.62, .79, .82, 1.71, 2.10

$\hat{E}_{10}$ detector: Decide YES iff $\hat{E}_{10} \leq 6$ (i.e., if six or fewer of the ten equiprobable cells are occupied).

Computations

To find the boundary points of the ten equiprobable intervals or cells one must solve

$$1 - \exp(-x^2) = p \text{ or } x = [-\ln(1-p)]^{\frac{1}{2}} \text{ for } p = .1, .2, \cdots, .9.$$

The cells and the number of PS-data points falling in them are as follows:

i	1	2	3	4	5	6	7	8	9	10
I	0, .325	.325, .472	.472, .597	.597, .715	.715, .833	.833, .957	.957, 1.097	1.097, 1.269	1.269, 1.517	1.517, ∞
f_i	1	1	0	0	2	0	0	1	1	15

Since four cells are empty, $\hat{E}_{10} = 6$ and the $\hat{E}_{10}$ detector decides YES.

In closing this section one should note

1. that the exact distribution of $\hat{E}$ and the limiting distribution of Sherman's $\tilde{S}_n$ are given by:

   If $F = F_0$, then

(a) $P(\hat{E}_{s+1} \leq r) = \sum_{i=1}^{r} \binom{s+1}{i} \sum_{j=0}^{i} (-1)^j \binom{i}{j} \left(\frac{i-j}{s+1}\right)^n$

and

(b) $\widetilde{S}_n$ is asymptotically $N(\mu_n, \sigma_n^2)$

where $\mu_n = [n/(n+1)]^{n+1}$ which tends to 1/e and

$\sigma_n^2 = [2n^{n+2} + n(n-1)^{n+2}]\,[(n+2)(n+1)^{n+2}]^{-1} - [n/(n+1)]^{2n+2}$

which is approximated by $(2e-5)/ne^2$;

2. that in view of the discreteness of the statistic, only a few significance levels can be obtained exactly;

3. that there is a need for a more extensive tabulation of the statistic $\widetilde{E}$; and

4. that there is a need for a "rule-of-thumb" for choosing the number s (of division points or $(s+1)$ of cells) for a given PS data sample size n in order to attain some preassigned goodness criteria.

In order to circumvent the difficulties above, one can make use of the Model I rank-sum detectors of the next section.

C. MODEL I RANK-SUM DETECTORS

The detectors of this section are based on statistics of the form

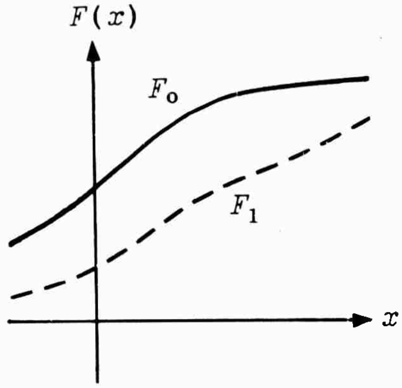
$$K(H) = \frac{1}{n} \sum_{i=1}^{n} H^{-1}[F_0(X_i)] \tag{24}$$

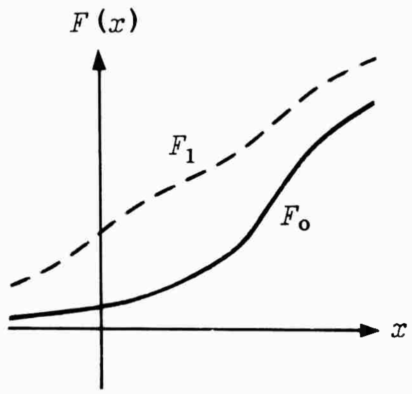
where H is a strictly increasing continuous cpf, and H^{-1} is its inverse. At first glance it is seen that $K(H)$ is the mean of the transformed observations $H^{-1}[F_0(X_1)], \cdots, H^{-1}[F_0(X_n)]$. In fact, if both H and F_0 are normal, exponential, or Rayleigh cpf's, then $K(H)$ is of the form $a\bar{X} + b$, i.e., is a linear function of the sample mean. (See table 5.)

Table 5.

H	$H^{-1}(u)$	F_0	$H^{-1}[F_0(X_i)]$	$K(H)$
$N(\mu_1, \sigma_1^2)$	$\sigma_1 \Phi^{-1}(u) + \mu_1$ where Φ is the $N(0,1)$ cpf.	$N(\mu_2, \sigma_2^2)$	$\left(\frac{\sigma_1}{\sigma_2}\right) X_i + \left(\mu_1 - \frac{\sigma_1 \mu_2}{\sigma_2}\right)$	$\left(\frac{\sigma_1}{\sigma_2}\right) \bar{X} + \left(\mu_1 - \frac{\sigma_1 \mu_2}{\sigma_2}\right)$
$1-e^{-x/\lambda_1}$ $(x \geq 0)$	$-\lambda_1 \ln(1-u)$	$1-e^{-x/\lambda_2}$ $(x \geq 0)$	$(\lambda_1/\lambda_2) X_i$	$(\lambda_1/\lambda_2) \bar{X}$
$1-e^{-x^2/\lambda_1}$ $(x \geq 0)$	$[-\lambda_1 \ln(1-u)]^{\frac{1}{2}}$	$1-e^{-x^2/\lambda_2}$ $(x \geq 0)$	$\sqrt{\lambda_1/\lambda_2}\ X_i$	$\sqrt{\lambda_1/\lambda_2}\ \bar{X}$

These statistics are, then, generalizations of the sample mean $\bar{X}$; and one might expect such statistics to have properties similar to the properties of $\bar{X}$ in the parametric cases even when there is no simple relation between H and F_0.





Consider the situation in which the pure-noise cpf F_o is above the noise-plus-signal cpf F_1, or equivalently, F_1 is to the right of F_o. In this case, if there is noise-plus-signal the PS data values $X_1, \cdots, X_n$ will "on the average" be larger than if there is only pure noise and one says that the noise-plus-signal values are stochastically larger. Consequently, if there is noise-plus-signal the $\{F_o(X_i)\}$ are stochastically larger as are the $\{H^{-1}[F_o(X_i)]\}$, and $K(H) = \frac{1}{n}\sum H^{-1}[F_o(X_i)]$. Hence, in the case outlined above, the $K(H)$ detector would decide YES iff $K(H)$ exceeded a level determined by n and α. Of course, the reverse decision rule would be employed by the detector in guarding against noise-plus-signal which is stochastically smaller.

In practice, the detectors are usually based on one of the following four statistics:

(a) $$\pi = -2 \sum_{i=1}^{n} \ln\,[F_o(X_i)]\,; \qquad (25)$$

(b) $$\pi' = -2 \sum_{i=1}^{n} \ln\,[1 - F_o(X_i)]\,; \qquad (26)$$

(c) $$\bar{U} = \frac{1}{n} \sum_{i=1}^{n} F_o(X_i); \text{ and} \tag{27}$$

(d) $$\bar{Z} = \frac{1}{n} \sum_{i=1}^{n} \Phi^{-1}[F_o(X_i)]\,. \tag{28}$$

Coincidentally, these four statistics result from choosing, as H, those distributions which have maximum Shannon information over (a') the right half-line; (b') the left half-line, (c') the unit interval and (d') the whole real line, respectively. Namely, (a'') the exponential $H(x) = 1 - e^{-x}\,(x \geq 0)$; ($b''$) the negative exponential, $H(x) = e^{x}$, $x < 0$; (c'') the standard uniform $H(x) = x$, $0 \leq x \leq 1$ and (d'') the standard normal cpf $H(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)\, dt$.

Thus, there is some intuitive justification for the choice of these four statistics, although they seem to be directed toward detection situations in which F_o and F_1 do not cross, i.e., in which the pure-noise distribution is stochastically larger or stochastically smaller than the noise-plus-signal distribution.

Example 6.

Pure-noise cpf: $F_o(x) = 1 - \exp(-.0625\, x^2)$ $(x \geq 0)$

PFA $\alpha = .005$

PS data: 3.2, 5.2, 1.2, 7.6, .8 $(n = 5)$

(a) π detector: Decide YES iff

$$-2 \sum_{i=1}^{5} \ln\,[F_o(X_i)] > 25.2$$ (the 99.5th percentile of the chi-square distribution with $2n = 10$ degrees of freedom).

(b) π' detector: Decide YES iff

$$-2 \sum_{i=1}^{5} \ln \left[1 - F_o(X_i)\right] < 2.16$$

(the 0.5th percentile of the chi-square distribution with $2n = 10$ degrees of freedom).

(c) $\bar{U}$ detector: Decide YES iff $(.2) \sum_{i=1}^{5} F_o(X_i) > .751$

(where .751 is the 99.5th percentile of the distribution of the mean of a sample of size 5 from a standard uniform distribution).

(d) $\bar{Z}$ detector: Decide YES iff $(.2) \sum_{i=1}^{5} \Phi^{-1}[F_o(X_i)] > 1.117$

(where 1.117 is the 99.5th percentile of the normal distribution with $\mu = 0$ and $\sigma^2 = .2$).

Computations

i	$X(i)$	$F_o(X(i)) =$ $1\text{-exp}\,[-.0625\, X^2(i)]$	$\ln F_o(X(i))$	$\Phi^{-1}[F_o(X(i))]$	$\ln\,[1-F_o(X(i))]$ $= -.0625\, X^2(i)$
1	.8	.040	-3.220	-1.750	- .040
2	1.2	.086	-2.450	-1.365	- .090
3	3.2	.457	- .782	- .108	- .640
4	5.2	.816	- .203	+ .900	-1.690
5	7.6	.973	- .027	+1.927	-3.610
		2.372	-6.682	- .396	-6.070

$$\sum F_o(X_i) = 2.372$$

$$\sum \ln F_o(X_i) = -6.682$$

$$\sum \Phi^{-1}[F_o(X_i)] = -.396$$

$$\sum \ln [1 - F_o(X_i)] = \sum [-.0625\ X^2(i)] = -6.070$$

(a) π detector

$$-2 \sum_{i=1}^{5} \ln [F_o(X_i)] = 13.364 \ngtr 25.2 \qquad \text{Decision: NO}$$

(b) π' detector

$$-2 \sum_{i=1}^{5} \ln [1 - F_o(X_i)] = 12.140 \nless 2.16 \quad \text{Decision: NO}$$

(c) $\bar{U}$ detector

$$(.2) \sum_{i=1}^{5} F_o(X_i) = .4744 \ngtr .750 \qquad \text{Decision: NO}$$

(d) $\bar{Z}$ dctector

$$(.2) \sum_{i=1}^{5} \Phi^{-1}[F_o(X_i)] = -.0792 \ngtr 1.117 \quad \text{Decision: NO}$$

Very recent results[12] indicate that the statistics of the form $\sum_{i=1}^{n} H^{-1}[F_o(X_i)]$ are really special cases of a more general class of statistics--the "likelihood-ratio" distribution-free statistics.

They are of the form

$$T_h = \sum_{i=1}^{n} \ln h'[F_0(X_i)] \tag{29}$$

where h is a differentiable strictly monotone cpf on the unit interval, i.e., h is monotone increasing and $h(0) = 0$, $h(1) = 1$, and of the form

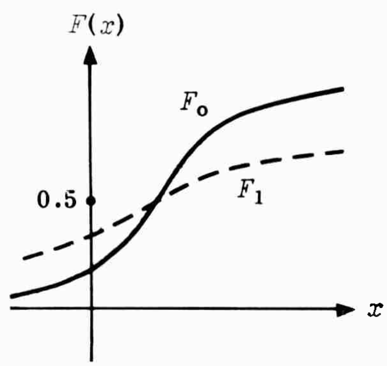
$$T' = \Sigma J[F_0(X_i)] \tag{30}$$

where $J'(u) = \frac{\partial}{\partial a} \ln h_a'(u) \Big|_{a=a_0}$, and $\{h_a\}$ is a family of differentiable, strictly increasing cpf's with $h_{a_0}(u) = u$.

Detectors based on this more general statistic reduce to the π', π, $\bar{U}$, and $\bar{Z}$ detectors when $h(u) = 1-(1-u)^a$; u^a; $ae^{bu}+c$ and $\Phi[\Phi^{-1}(u) - a]$, respectively. Of course, the number of detectors of this more general type is unlimited. However, the most common version other than the π, π', $\bar{U}$, and $\bar{Z}$ detectors, which intuitively seem good for the stochastically larger and stochastically smaller cases, is

$$\bar{Z}^2 = \sum_{i=1}^{n} \left\{\Phi^{-1}[F_0(X_i)]\right\}^2, \text{ which results for } h(u) = \Phi\left[\frac{\Phi^{-1}(u)}{a}\right]. \tag{31}$$

By intuition and/or an investigation[12] of the method of deriving this formula, one concludes that a $\bar{Z}^2$ detector has some optimal properties for situations in which the noise-plus-signal cpf F_1 and the pure-noise cpf F_0 have different dispersions or spreads. Exact methods of choosing h will be discussed in the section on goodness criteria. This section will be terminated with an example of the application of the $\bar{Z}^2$ detector.



*Example 7.**

Pure-noise cpf: F_o is $N(0,4)$

PFA $\alpha = .005$

PS data: -.472, -2.180, 6.684, 2.592, 3.332, 2.584
$n = 6$

$\bar{Z}^2$ detector: Decide YES iff $\sum_{i=1}^{n} \left\{\Phi^{-1}[F_o(X_i)]\right\}^2 > 18.55$

(the 99.5th percentile of the chi-square distribution with n degrees of freedom).

Computations

i	$X(i)$	$F_o(X(i))$	$\Phi^{-1}[F_o(X(i))]$	$\left\{\Phi^{-1}[F_o(X(i))]\right\}^2$
1	-2.180	.1379	-1.090	1.188
2	- .472	.4067	- .236	.056
3	2.584	.9018	1.292	1.669
4	2.592	.9025	1.296	1.680
5	3.332	.9521	1.666	2.776
6	6.684	.9996	3.342	11.169
				$18.538 = \bar{Z}^2$

Decision: NO

* In example 7, $\Phi^{-1}[F_o(x)]$ reduces to $x/2$. This type of simplification is not possible in general.

D. GOODNESS CRITERIA FOR MODEL I DETECTORS

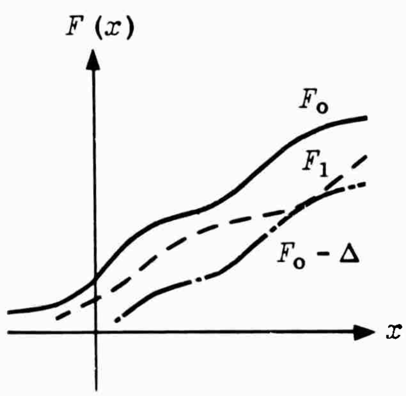
All statistical goodness criteria for detectors are based directly or indirectly on the PFA α and PFD β. Usually one fixes PFA α and is consequently led to criteria which (for fixed α) depend solely on the PFD β.

Assuming that both the pure-noise cpf F_o and the noise-plus-signal cpf F_1 are normal, one finds that for SDF detectors β depends only on σ_o/σ_1 and $(\mu_1 - \mu_o)/\sigma_1$, i.e., the ratio of the standard deviations and the ratio of the difference of means to σ_1. Hence, β can be completely tabulated subject only to limitations of computation time and costs.

The detectors being considered here are all SDF and, hence, β depends only on $F_o F_1^{-1} = (F_1 F_o^{-1})^{-1}$. However, even with this significant reduction of the possible alternatives, it is impossible to tabulate or give formulae for all situations of interest. Consequently, a statement of the form "The T_1 detector is better than the T_2 detector" cannot be made without some rather severe qualifications.

In order to make some recommendations as to the relative merits of Model I detectors, one must restrict consideration to certain special classes of alternatives, as follows.

1. One-sided Bands and max-min PFD β



Based on the statistical work of Chapman[9] one considers for each strictly increasing continuous pure noise cpf F_o the class $\mathcal{G}(F_o, \Delta)$ of all noise-plus-signal cpf's F_1 such that

(a) $F_o > F_1$ and

(b) $\max\limits_x [F_o(x) - F_1(x)] = \Delta$,

where $0 < \Delta < 1$.

For the class $\mathcal{G}(F_0, \Delta)$, one says a T_1 detector is max-min PFD better than a T_2 detector if $\bar{\beta}(T_1, \Delta) < \bar{\beta}(T_2, \Delta)$ and $\underline{\beta}(T_1, \Delta) < \underline{\beta}(T_2, \Delta)$ where $\bar{\beta}$ and $\underline{\beta}$ are, respectively, the supremum and infimum of the indicated detectors for the class $\mathcal{G}(F_0, \Delta)$.

The following curves contain slots of $\bar{\beta}$ vs. Δ and $\underline{\beta}$ vs. Δ for the Model I detectors based on the statistics π, π', D_n^-, $n\omega_n^2$, $\bar{U}$, $Q'_{.2}$, $Q'_{.5}$, and $Q'_{.9}$.

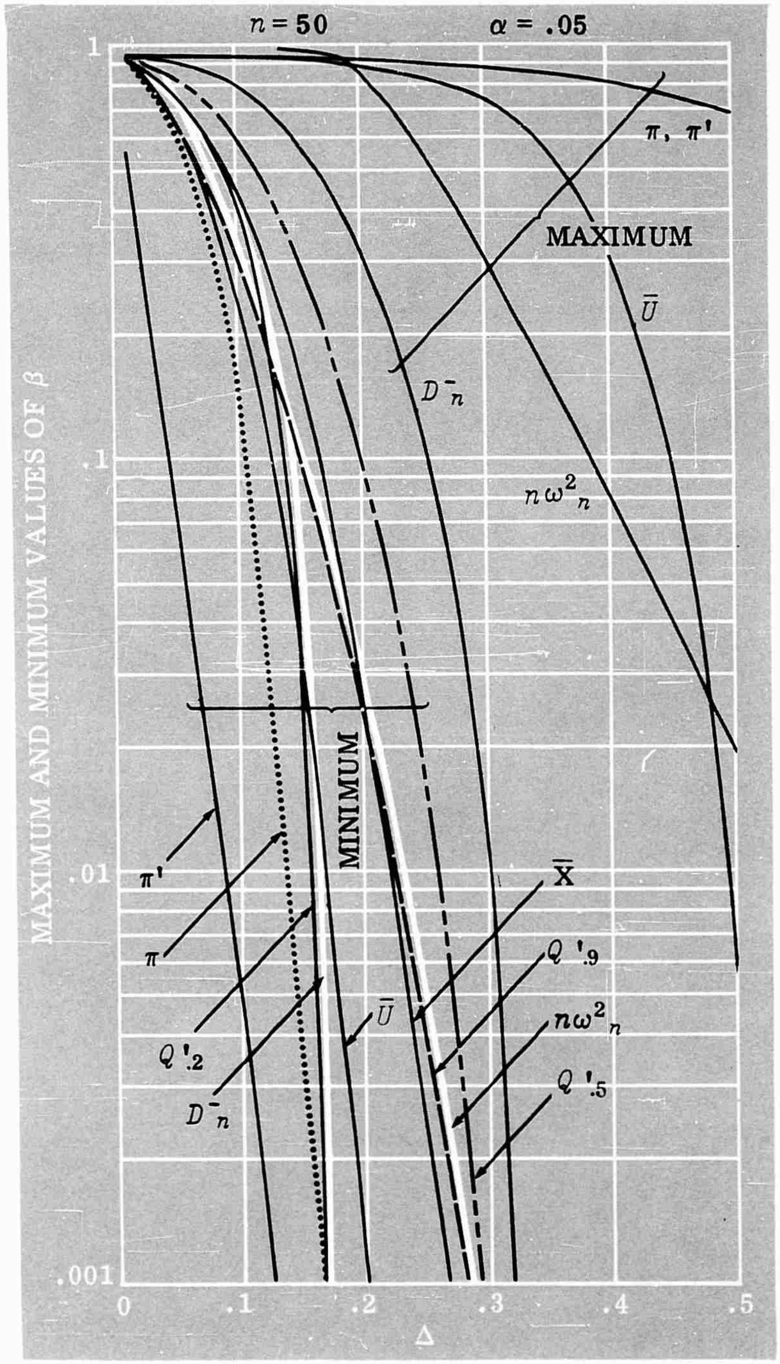
The maximum PFD's, the $\bar{\beta}$'s, for the statistics Q'_p are all $\bar{\beta}(Q'_p, \Delta) = 1 - \alpha$ and are not plotted. Also, on these graphs, which are plotted for $\alpha = .01$ and $.05$, and $n = 50$ and 100, there is a single graph labelled "$\bar{X}$," which is a plot of PFD β vs. Δ for the parametric case in which both the pure-noise cpf F_0 and the noise-plus-signal cpf F_1 are normal cpf's with equal variances and for which (a) and (b) above are satisfied. All the other graphs are completely DF, in the sense that they are valid for any strictly increasing continuous cpf F_0 and its alternative class $\mathcal{G}(F_0, \Delta)$.

From these graphs one concludes that

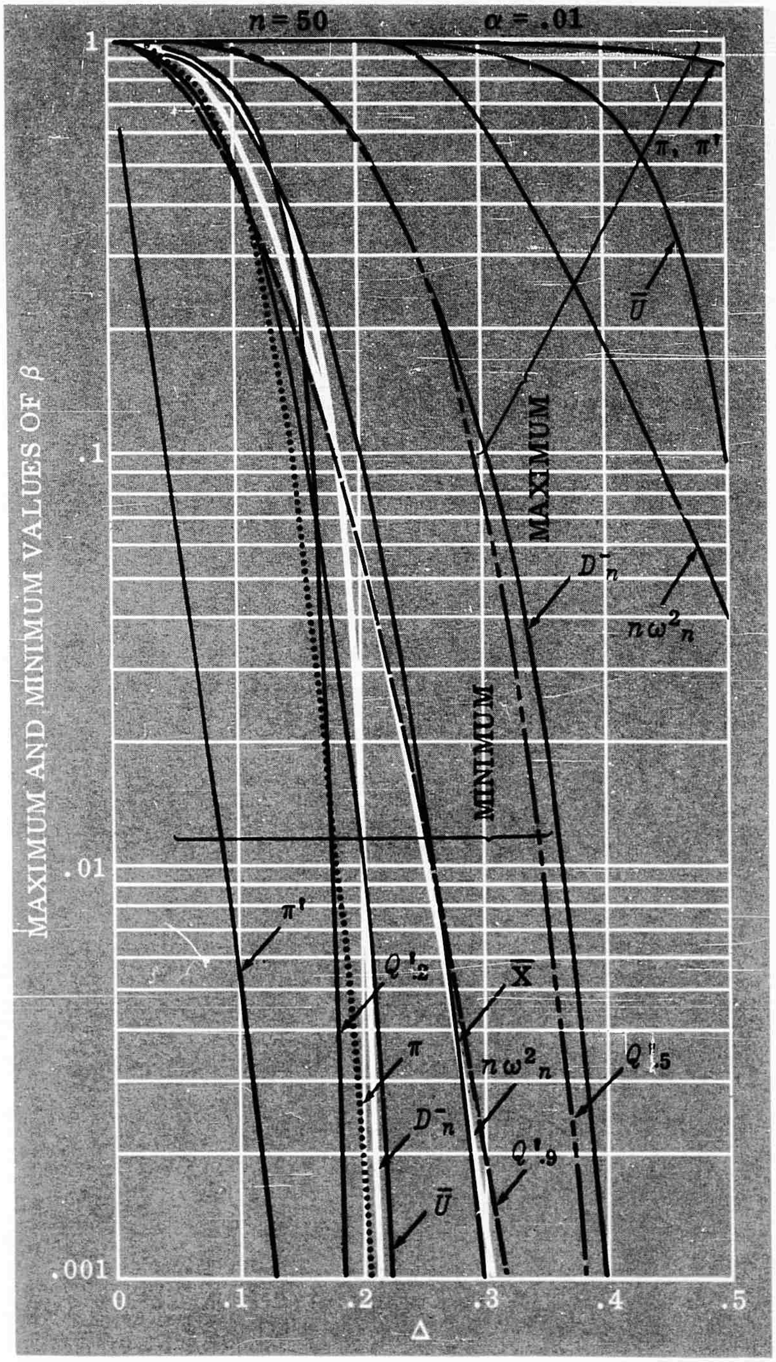
1. there is a general tendency for detectors with high maximum PFD's to have low minimum PFD's, and vice versa; hence,

2. no one of the Model I detectors based on π, π', D_n^-, $n\omega_n^2$, $\bar{U}$, $Q'_{.2}$, $Q'_{.5}$, $Q'_{.9}$ is max-min better than any other one for all of the cases considered.

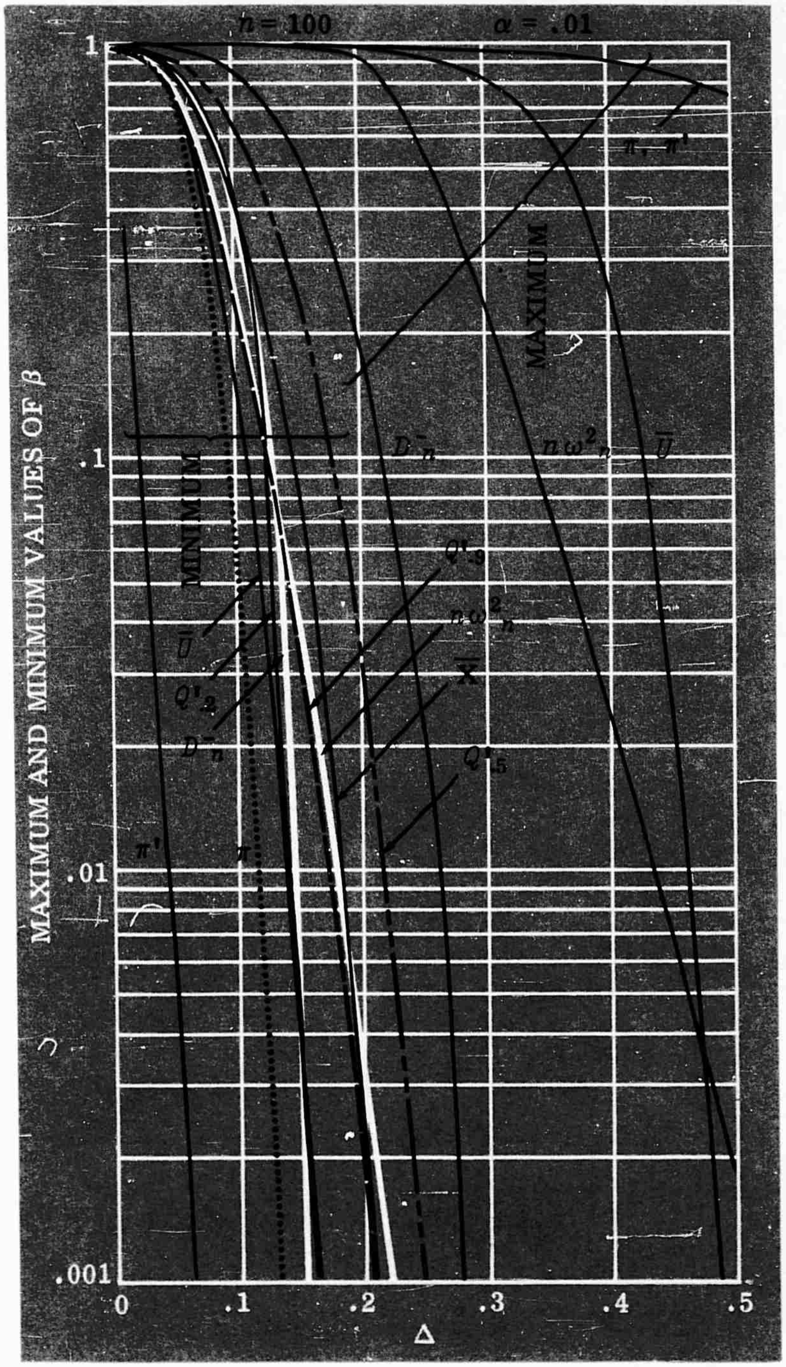
This result is somewhat discouraging to one who wishes to order the detectors as to goodness. It is further discouraging to note that as of this writing there are not available max-min PFD data for the other Model I detectors being studied, D_n, W_n^2, $\hat{Q}$, D_n^-, $\tilde{Q}_s$, $\bar{Z}$, $\bar{Z}^2$, etc. One therefore seeks other criteria.



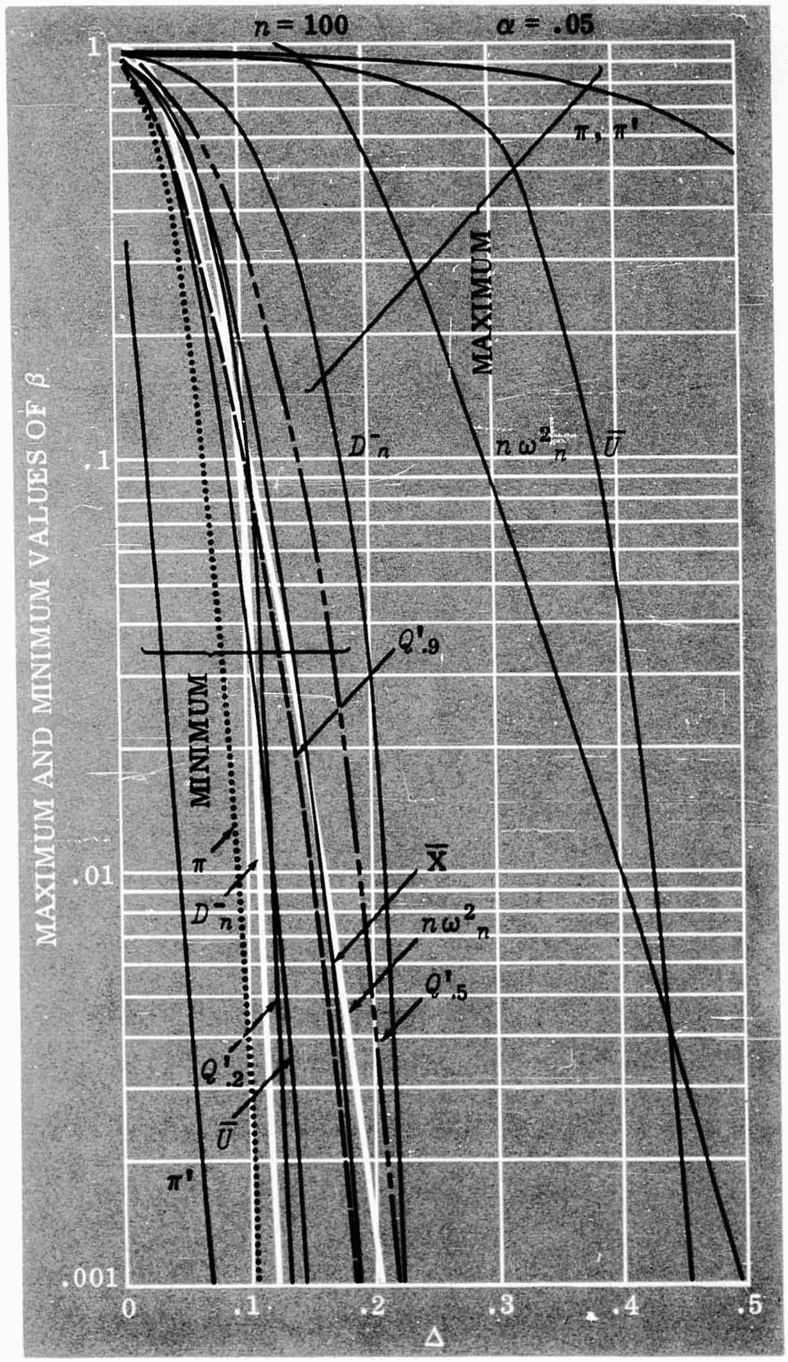
Maximum and minimum PFD for one-sided noise.



Maximum and minimum PFD for one-sided naise. (Continued)

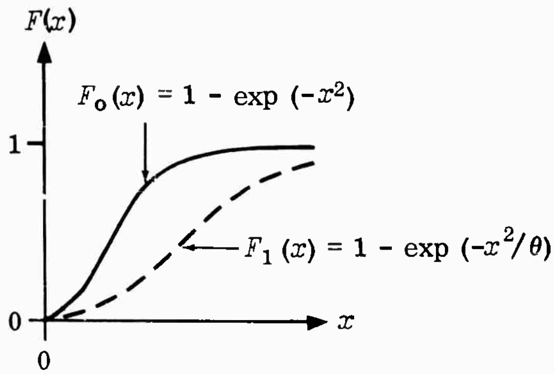


Maximum and minimum PFD for one-sided noise. (Continued)



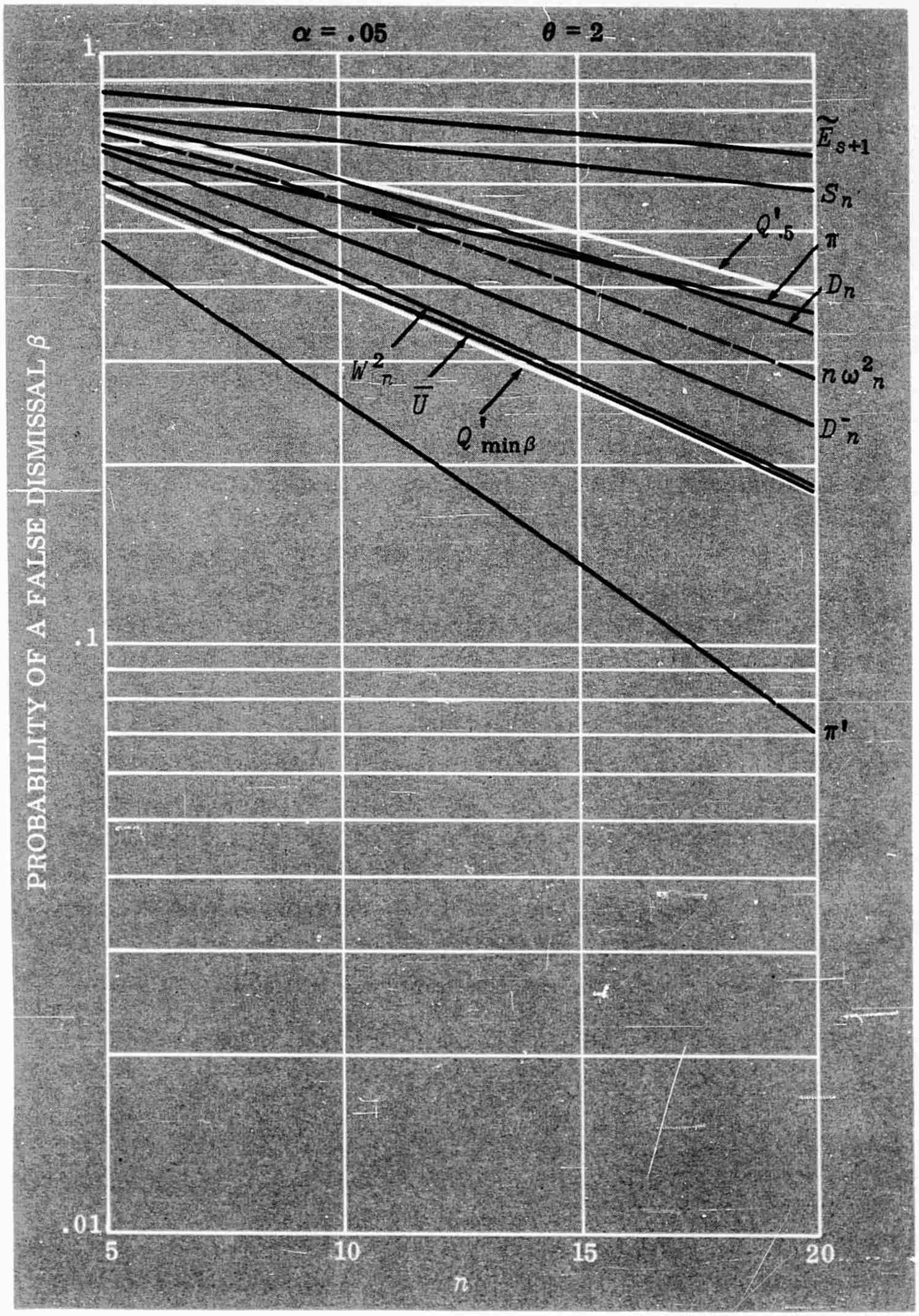
Maximum and minimum PFD for one-sided noise. (Continued)

2. PFD for Rayleigh Alternatives

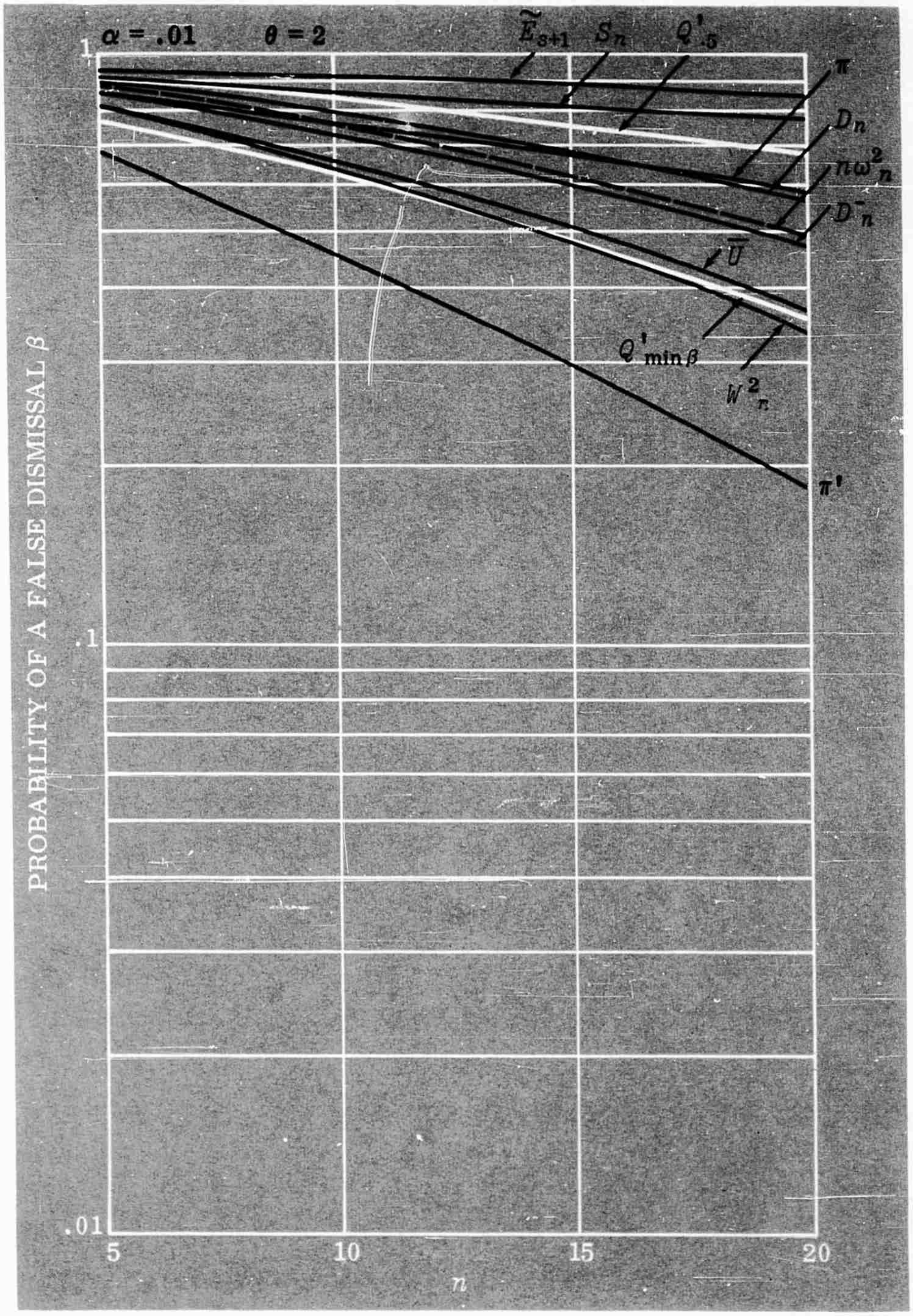


Some engineers have found that Rayleigh-type distributions occur frequently in practice. Consequently, it is not unreasonable to consider the performance of Model I detectors in the situation for which $F_0(x) = 1 - \exp(-x^2)$ and $F_1(x) = 1 - \exp(-x^2/\theta)$, as illustrated. (A more accurate plot of F_0 is given on page 30.)

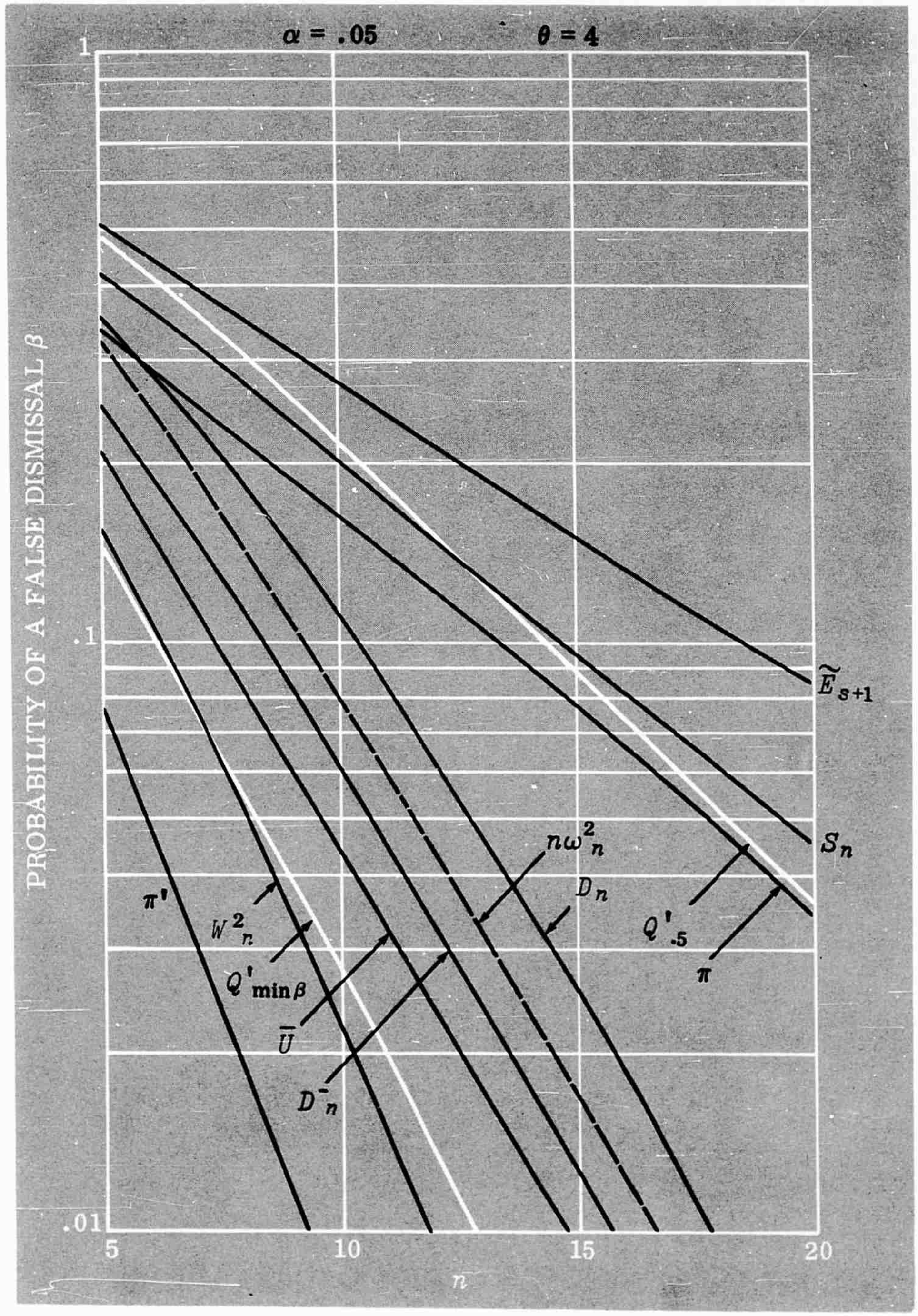
Since the mathematical expressions for exact PFD are almost completely intractable, one has to resort to Monte Carlo methods (10,000 trials) to obtain accurate estimates of the PFD's except for the Q' detectors in the plots of β vs. n for fixed values θ, where only the values for which $\theta = 2$ and 4, and $n = 5$, 10, 15, and 20 are computed. The curve labeled $Q'_{\min\beta}$ represents the value of β for the optimal choice of the quantile. Further, one might intuitively divide the graphs into two sets--one for those detectors which are based on the "one-sided" statistics D^-_n, π, π' and $\bar{U}$ and the other for those detectors based on the "two-sided" statistics D_n, $n\omega^2_n$, S_n, W^2_n, and $\widetilde{E}_{s+1}$. This latter division of the class of detectors was made because, prior to making the computations, it seemed intuitively clear that the "one-sided" detectors should have, in general, better performances against the one-sided alternatives $1 - \exp(-x^2/2)$ and $1 - \exp(-x^2/4)$. This suspicion is, however, not borne out by the plots of the Monte Carlo computations.



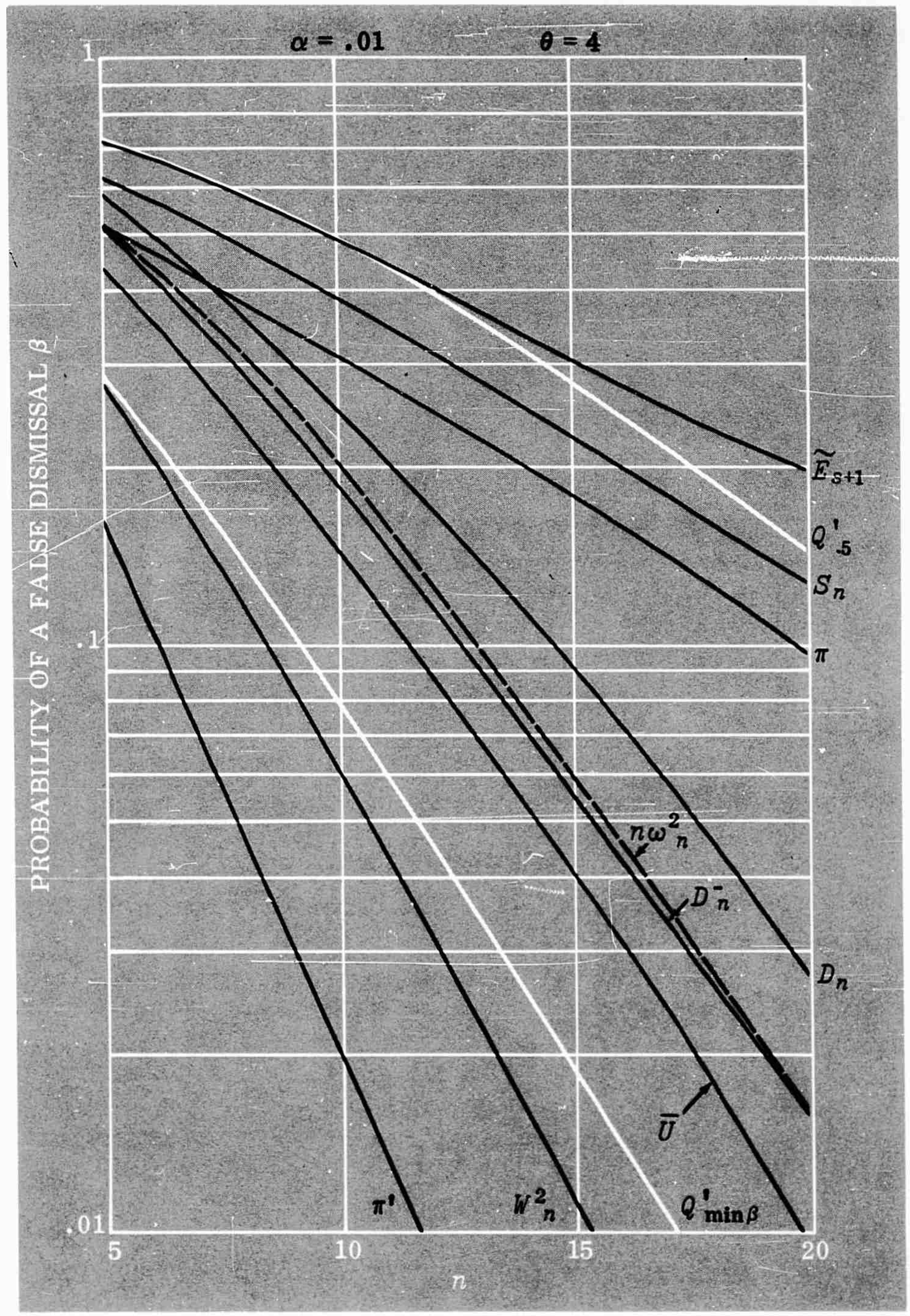
PFD for Rayleigh noise.



PFD for Rayleigh noise. (Continued)



PFD for Royleigh noise. (Continued)



PFD for Royleigh noise. (Continued)

From these plots one can make the following general statements in which "$\}$" indicates "generally better than" for the cases considered.

(a) For the "two-sided" detectors

$$W^2_n \;\}\; n\omega^2_n \;\}\; D_n \;\}\; S_n \;\}\; \widetilde{E}_{s+1}$$

(b) For the "one-sided" detectors

$$\pi' \;\}\; \bar{U} \;\}\; D^-_n \;\}\; \pi \;\}\; Q'_{.5}$$

(c) There is a significant overlap of the ranges of PFD for the two sets of detectors.

(d) Of the detectors studied for the cases considered, π' is clearly the best and E_{s+1} is clearly the worst.

These results, of course, depend very heavily on the class of alternatives considered, and the strongest statement one can make on the basis of these results is the following.

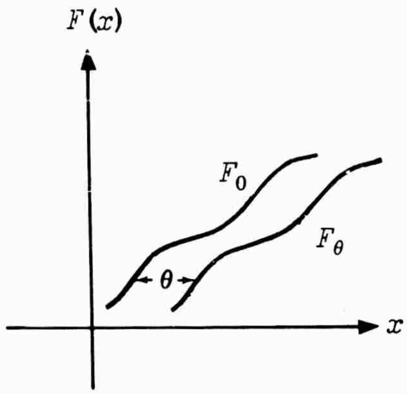
(e) <u>If one wishes to use a Model I detector which (besides its DF property) performs well against Rayleigh alternatives, one should choose a π' detector and avoid using an E_{s+1} detector</u>.

In order to make comparable statements concerning other classes of pure noise and noise-plus-signal cpf's one would have to make comparable computations. Of course, the amount of such computation possible is only limited by time and money considerations.

3. Asymptotic Relative Efficiency (ARE)

The next comparison of Model I detectors will be made on the basis of ARE. One recalls that, roughly, the ARE, $A(T_1, T_2)$, of a T_1 detector with respect to a T_2 detector is the limit, $\lim \frac{n(T_2)}{n(T_1)}$, of the ratio of sample sizes for fixed α and β in the

presence of increasingly weak signals. Of course, as the signal becomes increasingly weak, the noise-plus-signal cpf F_1 tends to F_o; and, as one might suspect, the limit, $A(T_1, T_2)$, is sensitive to the form of the F_o and the manner in which the F_1's tend to F_o.



$F_\theta(x) = F_o(x-\theta)$

Another consideration of great practical importance is the fact that in a large number of cases of interest the derivation of a computable expression for ARE is almost impossible. This situation is reflected in table 6, which treats only detectors based on $\overline{Z}$, $\overline{U}$, π, and π', translation alternatives (as illustrated here), and certain classes of pure-noise cpf's F_o. Several numerical entries are missing because they have not yet been computed.

Table 6. ARE's for Model I Detectors

Pure-noise cpf $F_o(x)$	Noise-plus-signal F_θ where $\theta \searrow 0$	$A(\overline{Z}, \overline{U})$	$A(\overline{Z}, \pi)$	$A(\overline{Z}, \pi')$	$A(\pi, \overline{U})$	$A(\pi', \overline{U})$	$A(\pi, \pi')$
$\Phi(x)$ (Normal)	$\Phi(x-\theta)$	1.05					
U_o (Uniform)	$F_\theta(x) = x-\theta$, $0 \le x \le 1+\theta$	∞			∞	∞	1.00
H_e (Exponential)	$F_\theta(x) = 1-e^{-x+\theta}$, $\theta \le x$	∞	∞		.33		
e^x, $x \le 0$ (Negative exponential)	$e^{x-\theta}$					.33	
$(1+e^{-x})^{-1}$ (Logistic)	$(1+e^{-x+\theta})^{-1}$	.955	1.27	1.27	.75	.75	1.00
$\begin{cases} e^{x}/2, & x \le \theta \\ 1-e^{-x}/2, & x > \theta \end{cases}$ (Double exponential)	$\begin{cases} e^{x-\theta}/2, & x \le \theta \\ 1-e^{-x+\theta}/2, & x > \theta \end{cases}$	.85					
$\frac{1}{2} + \frac{1}{\pi}\tan^{-1} x$ (Cauchy)	$\frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x-\theta)$	.71					

From table 6 it may be concluded:

(a) that there is no clear-cut ordering of the detectors on the basis of known ARE's; but

(b) that if one wishes good asymptotic performance (in the presence of increasingly weak signals) for logistic type distributions one should use a $\bar{U}$ detector and avoid π and π' detectors.

The final goodness criteria to be discussed are adaptations of very recent statistical results of Bell and Doksum[12].

4. Uniformly Minimum PFD and Almost Locally Minimum PFD Detectors

In keeping with the ideas developed in the preceding section, two other optimal properties, which depend on the particular class of alternatives, will be discussed and illustrated here. UM PFD and ALM PFD detectors will be considered here to be as defined in the Introduction.

One recalls that for SDF Model I detectors, β is a function solely of $h_\theta = F_\theta F_0^{-1}$, where F_0 is the pure-noise cpf and F_θ is the noise-plus-signal cpf. By restricting attention to classes $\{F_\theta\}$ which satisfy

(1) $\lim_{\theta \to \theta_0} h_\theta(u) = u$ for all u; and

(2) $\left.\dfrac{\partial}{\partial \theta} \ln h_\theta'(u)\right|_{\theta=\theta_0} = J_0(u)$ exists;

one can make use of the following theorem of Bell and Doksum:[12]

Theorem 6. An SDF detector whose YES-NO decisions are based on the statistic

$$T' = \sum_{i=1}^{n} J_0\left[F_0(X_i)\right]$$

(i) is a UM PFD detector for each class of noise-plus-signal cpf's with $h_\theta'(u) = \exp\ [a(\theta)\ J_0(u) + b(\theta)]$ where $a(\theta) \neq 0$ for $\theta \neq \theta_0$ and $a(\theta_0) = b(\theta_0) = 0$; and

(ii) is an ALM PFD detector for each class of noise-plus-signal cpf's with $h_\theta'(u) = \exp\ [a(\theta)\ J_0(u) + b(\theta) + Q(u, \theta)]$ where $a(\theta)$ and $b(\theta)$ are as in (i), and $\lim Q(u, \theta)/(\theta - \theta_0) = 0$ for all u, where the limit is taken as θ tends to θ_0.

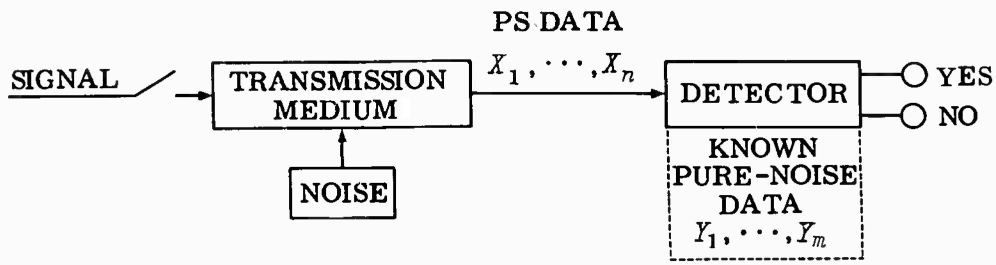
Employing this theorem and the fact that a detector based on a statistic T' is equivalent to that based on the statistic $c(\theta)T' + d(\theta)$, where $c(\theta) \neq 0$, one can construct table 7, which illustrates and summarizes known results in this direction. The complexity of the formulas and the vacant spaces in the table indicate the depth of the problem here. (Note that each detector is an SDF detector, which also has certain optimal properties for specified classes of noise-plus-signal cpf's.)

Table 7. Optimal Model I Detectors for Certain Noise Classes $\{F_\theta\}$
$\left[h_\theta = F_\theta F_o^{-1} \text{ and } \lim\limits_{\theta \to \theta_o} h_\theta(u) \equiv u\right]$

$h_\theta(u)$	θ_o	Statistic for UM PFD Detector	Statistic for ALM PFD Detector	General Example	Specific Example
$\exp[a(\theta)J(u)+b(\theta)]$ $= h_\theta'(u)$	θ_o	$\Sigma\, J[F_o(X_i)]$	$\Sigma\, J[F_o(X_i)]$	$F_o(x)$ vs. $\int h_\theta[F_o(x)]\, F_o'(x)\, dx$	
$\exp[a(\theta)J(u)+b(\theta) + Q(u,\theta)]$ $= h_\theta'(u)$ where $Q(u,\theta) = 0(\theta-\theta_o)$	θ_o	$\Sigma\, J(F_o(X_i)) + c(\theta)\, \Sigma\, Q[F_o(X_i), \theta]$	$\Sigma\, J[F_o(X_i)]$	$F_o(x)$ vs. $\int h_\theta(F_o(x))\, F_o'(x)\, dx$	
u^θ	1	$\pi = -2\, \Sigma \ln F_o(X_i)$	π	$F_o(x)$ vs. $F_o^\theta(x)$	
$1-(1-u)^\theta$	1	$\pi' = -2\, \Sigma \ln[1-F_o(X_i)]$	π'	$F_o(x)$ vs. $[1-(1-F_o)^\theta]$	$1-e^{-x}$ vs. $1-e^{-x\theta}$
$(1-\theta)u+\theta u^2$	0	$\Sigma \ln[1-\theta+2\theta\, F_o(X_i)]$	$\overline{U} = \frac{1}{n}\Sigma\, F_o(X_i)$	$F_o(x)$ vs. $[(1-\theta)F_o(x)+\theta\, F_o(x)^2]$	
$(1-\theta)u+\theta u^{K+1}$	0	$\Sigma \ln[1-\theta+(K+1)\theta F_o^{K+1}(X_i)]$	$\Sigma\, [F_o(X_i)]$[illegible]	$F_o(x)$ vs. $[(1-\theta)F_o(x)+\theta\, F_o^{K+1}(x)]$	
$\frac{e^{\theta u}-1}{e^\theta -1}$	0	$\overline{U}$	$\overline{U}$	$F_o(x)$ vs. $\frac{e^{\theta F_o(x)}-1}{e^\theta -1}$	
$\frac{u}{u+e^\theta(1-u)}$	0	$\Sigma \ln\{F_o(X_i) + e^\theta[1-F_o(X_i)]\}$	$\overline{U}$	$F_o(x)$ vs. $\frac{F_o(x)}{F_o(x)+e^\theta[1-F_o(x)]}$	$\frac{1}{1+e^{-x}}$ vs. $\frac{1}{1+e^{-x+\theta}}$
$G[G^{-1}(u)-\theta]$	0	$\Sigma \ln(G'\{G^{-1}[F_o(X_i)-\theta]\}) - \Sigma \ln(G'\{G^{-1}[F_o(X_i)]\})$	$\Sigma\, \frac{G''\{G^{-1}[F_o(X_i)]\}}{G'\{G^{-1}[F_o(X_i)]\}}$	$F_o(x)$ vs. $F_o(x-\theta)$	above; and $\Phi(x)$ vs. $\Phi(x-\theta)$
$G\left[\frac{G^{-1}(u)}{\theta}\right]$	1	$\Sigma \ln\left(G'\left\{\frac{G^{-1}[F_o(X_i)]}{\theta}\right\}\right) - \Sigma \ln(G'\{G^{-1}[F_o(X_i)]\})$	$\Sigma\, \frac{G''\{G^{-1}[F_o(X_i)]\, G^{-1}[F_o(X_i)]\}}{G'\{G^{-1}[F_o(X_i)]\}}$	$F_o(x)$ vs. $F_o(x/\theta)$	$\Phi(x)$ vs. $\Phi(x/\theta)$

Table 8. Statistical Distributions for Model I Detectors

Statistic	Distribution	Source of Tables, Formulas, etc. (Numbers refer to Bibliography.)
D_n	(1-sample Kolmogorov)	21, p. 423 ff
D^+_n, D^-_n	(1-sample one-sided Kolmogorov)	22
D'_n		23
$n\omega^2_n$	(1-sample Cramér-von Mises)	20; 21, p. 443 ff
W^+_n, W^-_n	(Equivalent to $\bar{U}$)	
W^2_n	(Anderson-Darling)	24
Q'_p	Binomial (sign)	21, p. 362 ff
$\tilde{Q}$, $\tilde{Q}_s$	Asymptotically chi-square	21, p. 49 ff; 8, p. 421 ff
$\hat{Q}$, $\hat{Q}_s$	Asymptotically chi-square (Matusita)	25; 21, p. 49 ff
$\tilde{S}_n$	Asymptotically normal (Sherman)	26; 21, p. 477 ff
S'_n	(Kimball-Moran)	
$\hat{E}_{s+1}$	(Empty cell)	21, p. 454 ff
π, π'	Chi-square	9; 21, p. 49 ff
$\bar{U}$	Sample mean of rectangular; asymptotically $N\left(\frac{1}{2}, \frac{n}{12}\right)$	27, p. 257 11
$\bar{Z}$	$N(0, 1)$	21, p. 1 ff
$\bar{Z}^2$	Chi-square	12; 21, p. 49 ff



COMPARE $X_1, \cdots, X_n$ AND $Y_1, \cdots, Y_m$

III. MODEL II DETECTORS

The Model II detector is, as would be expected, an extension of the Model I detector. In Model II PS data $X_1, \cdots, X_n$ are available, as well as known pure-noise data $Y_1, \cdots, Y_m$. The decision-making process then consists of comparing $X_1, \cdots, X_n$ with $Y_1, \cdots, Y_m$ and deciding YES iff they are "too far" apart in some sense. Further, one recalls that by imposing the reasonable SDF and Scheffé conditions of theorem 2, page 10, one is led to the class of rank detectors which base their YES-NO decisions solely on the ranks $R(X_1), \cdots, R(X_n), R(Y_1), \cdots, R(Y_m)$ of the X's and Y's in the combined sample $X_1, \cdots, X_n, Y_1, \cdots, Y_m$.

As in the case of Model I detectors one considers three "natural" subclasses of Model II detectors:

1. <u>Sample cpf (SDF Scheffé Model II) detectors</u>, which employ statistics based on differences of the PS-data cpf F_n and the pure-noise-data cpf G_m.

2. <u>Run-block (SDF Scheffé Model II) detectors</u>, which base their decisions on the number of PS-data values $X_1, \cdots, X_n$ which fall in certain intervals determined by certain (preassigned) order statistics of the pure-noise sample $Y_1, \cdots, Y_m$.

3. <u>Rank-sum (SDF Scheffé Model II) detectors</u>, whose decisions are based on sums of functions of the ranks $R(X_1), \cdots, R(X_n), R(Y_1), \cdots, R(Y_m)$.

In the succeeding three sections general formulas and specific examples of the common types of Model II detectors are developed. The fourth of the succeeding sections introduces a detector based on a randomized statistic, which is relatively new. This new class of detector introduces artificial noise into the detector in order to circumvent some difficulties arising in the use of the more classical DF detectors.

In the latter sections of this chapter, the various Model II detectors are compared in terms of such previously developed goodness criteria as PFA α, PFD β, and max-min β.

In describing the Model II detectors it is necessary to recall some notation and to introduce other notation.

1. PS sample: $X_1, \cdots, X_n$ Order statistics: $X(1) < \cdots < X(n)$

2. Pure-noise sample: $Y_1, \cdots, Y_m$ Order statistics: $Y(1) < \cdots < Y(m)$

3. Combined sample: $\{Z_1, \cdots, Z_N\} = \{X_1, \cdots, Y_n; Y_1, \cdots, Y_m\}$ where $N = n + m$

4. Combined sample order statistics: $Z(1) < \cdots < Z(N)$

5. Degenerate cpf: ϵ where $\epsilon(x) = 0$ or 1 according as $x < 0$ or $x \geq 0$

6. PS sample cpf: F_n where $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \epsilon(x - X_i)$

7. Pure-noise sample cpf: G_m where $G_m(y) = \frac{1}{m} \sum_{j=1}^{m} \epsilon(y - Y_j)$

8. Combined sample cpf: $\bar{F}$ where $\bar{F}(z) = \frac{1}{N} \sum_{r=1}^{N} \epsilon(z - Z_r)$

9. Rank of X_i in combined sample: $R(X_i)$

   Rank of Y_j in combined sample: $R(Y_j)$

Some of the more useful relations between these quantities are:

1. $R(X_i) = \sum_{r=1}^{n} \epsilon(X_i - Z_r) = N\bar{F}(X_i)$

2. $\bar{F}(z) = \dfrac{nF_n(z) + mG_m(z)}{n+m}$

3. $\sum_{i=1}^{n} R(X_i) + \sum_{j=1}^{m} R(Y_j) = \sum_{r=1}^{N} r = \frac{1}{2} N(N+1)$

4. The possible ranks are $\{1, 2, \cdots, N\}$ and, consequently, the collection $\{R(X_1), \cdots, R(X_n)\}$ of X-ranks uniquely determines the collection $\{R(Y_1), \cdots, R(Y_n)\}$ of Y-ranks and vice versa.

A. MODEL II SAMPLE Cpf DETECTORS

The detectors of this type are essentially extensions of the Model I sample cpf detectors.

1. Kolmogorov-Smirnov Detector

$$D(n, m; r; \Psi) = \sup_x \left\{ \left| F_n(x) - G_m(x) \right| \sqrt{\Psi[\bar{F}(x)]} \right\}^r \tag{32}$$

The usual variations of this statistic are

$$D(n, m) = \sup_x |F_n(x) - G_m(x)| \tag{33}$$

$$= \max_r |F_n(Z(r)) - G_m(Z(r))| = m^{-1} \max |NF_n(Z(r) - r|$$

$$= \max \left[\max_i \left| \frac{i}{n} - G_m(X(i)) \right|, \max_j \left| F_n(Y(j)) - \frac{j}{m} \right| \right]$$

$$D^+(n, m) = \sup_x [F_n(x) - G_m(x)] = \max_r [F_n(Z(r)) - G_m(Z(r))] \tag{34}$$

$$= \max \left\{ \max_i \left[\frac{i}{n} - G_m(X(i)) \right], \max_j \left[F_n(Y(j)) - \frac{j}{m} \right] \right\}$$

$$D^-(n, m) = \sup_x [G_m(x) - F_n(x)] = \max_r [G_m(Z(r)) - F_n(Z(r))] \tag{35}$$

$$= \max \left\{ \max_i \left[G_m(X(i)) - \frac{i}{n} \right], \max_j \left[\frac{j}{m} - F_n(Y(j)) \right] \right\}$$

2. Cramér-von Mises Detector

Corresponding to the two-sample Kolmogorov-Smirnov detector is the two-sample Cramér-von Mises detector which bases its decisions on statistics of the form below.

$$W(n, m; r, \Psi) = \int_{-\infty}^{\infty} \left\{ |F_n(x) - G_m(x)| \sqrt{\Psi[\bar{F}(x)]} \right\}^r d\bar{F}(x) \tag{36}$$

The common versions of this statistic are

$$W^2(n,m) = \int_{-\infty}^{\infty} [F_n(x) - G_m(x)]^2 \, d\,\bar{F}(x) \tag{37}$$

$$= \frac{1}{N} \sum_{r=1}^{N-1} [F_n(Z(r)) - G_m(Z(r))]^2$$

$$= (1/Nm^2n^2) \sum_{r=1}^{N-1} [NnF_n(Z(r)) - r\ n]^2 \ ;$$

$$W^+(n,m) = \int_{-\infty}^{\infty} [F_n(x) - G_m(x)] \, d\,\bar{F}(x); \text{ and} \tag{38}$$

$$W^-(n,m) = \int_{-\infty}^{\infty} [G_m(x) - F_n(x)] \, d\,\bar{F}(x). \tag{39}$$

The two one-sided versions are related by $W^+(n,m) = -W^-(n,m) = \frac{n+1}{2m} + \frac{1}{2} - \sum_{i=1}^{n} R(X_i)$, which is equivalent to the Mann-Whitney-Wilcoxon statistic to be treated later.

For an illustration of the computations involved with Kolmogorov-Smirnov and Cramér-von Mises detectors, one can consider the following example.

Example 8.

PFA $\alpha = .005$ $\quad n = 5,\ m = 6$

$D(n,m)$ detector: Decide YES iff

$$\max_r |F_n(Z(r)) - G_m(Z(r))| > .833^*$$

$W^2(n,m)$ detector: Decide YES iff

$$\frac{1}{N}\sum_{r=1}^{N-1} [F_n(Z(r)) - G_m(Z(r))]^2 > .766^*$$

PS data (X_i): 3.2, -1.4, 7.0, 4.5, 0.7

Pure-noise data (Y_j): -1.6, 5.4, -0.4, -0.3, 2.5, 1.0

r	$Z(r)$	Sample	$F_n(Z(r))$	$G_m(Z(r))$	$F_n(Z(r)) - G_m(Z(r))$	$[F_n(Z(r)) - G_m(Z(r))]^2$
1	-1.6	Y	0	.167	-.167	.028
2	-1.4	X	.200	.167	.033	.001
3	-0.4	Y	.200	.333	-.133	.018
4	-0.3	Y	.200	.500	-.300	.090
5	0.7	X	.400	.500	-.100	.010
6	1.0	Y	.400	.667	-.267	.071
7	2.5	Y	.400	.833	-.433	.187
8	3.2	X	.600	.833	-.233	.054
9	4.5	X	.800	.833	-.033	.001
10	5.4	Y	.800	1.000	-.200	.040
11	7.0	X	1.000	1.000	.000	.000
						.500

$D(n,m)$ detector: $D(n,m) = .433 < .833$ Decision: NO

$W^2(n,m)$ detector: $W^2(n,m) = \frac{1}{11}(.500) = .045 < .766$ Decision: NO

* Obtained from ref. 21, p. 443-444. Since the exact value PFA $\alpha = .005$ was not tabulated, ".766" is an interpolated value, and ".833" yields $\alpha = .004$.

From the table above it is easy to compute the values of some of the other statistics of importance, such as

$D^{+}(n, m) = .033$; $D^{-}(n, m) = .433$ and

$\Sigma R(X) = 2 + 5 + 8 + 9 + 11 = 35.$

3. Sign-Quantile Detector

The final Model II sample cpf detector of importance is the sign-quantile detector. The rationale for its use is the immediate extension of that for the Model I sign-quantile detector, which is, roughly: if the unknown continuous pure-noise cpf F_0 and the unknown continuous PS-data cpf F_1 are equal, then one would expect the PS sample $X_1, \cdots, X_n$ and the known pure-noise sample $Y_1, \cdots, Y_m$ to be "well interspersed." More specifically, it would be expected that approximately np of the X's and mp of the Y's are less than or equal to the 100 p^{th} percentile $Z(Np)$ of the combined sample $Z_1, \cdots, Z_N$.

Consequently, for $0 = q_0 < q_1 < q_2 < \cdots < q_s < q_{s+1} = 1$

$$\left[F_n(Z(Nq_i)) - F_n(Z(Nq_{i-1})) - (q_i - q_{i-1})\right]$$

and

$$\left[G_m(Z(Nq_i)) - G_m(Z(Nq_{i-1})) - (q_i - q_{i-1})\right]$$

are both measures of deviation from the "well-interspersed" situation. Customarily, one uses the notation $p_i = q_i - q_{i-1}$;

$$n_{1i} = n\left[F_n(Z(Nq_i)) - F_n(Z(Nq_{i-1}))\right]$$

and

$$n_{2i} = m\left[G_m(Z(Nq_i)) - G_m(Z(Nq_{i-1}))\right].$$

Hence

$$\left[\frac{n_{1i}}{n} - p_i\right] \text{ and } \left[\frac{n_{2i}}{m} - p_i\right]$$

are the above-mentioned measures of dispersion.

For visualization it is perhaps helpful to consider table 9.

Table 9

Interval \ Sample	PS Data Frequency	Pure-Noise Data Frequency	TOTALS
$(-\infty, Z(Nq_1)]$	n_{11}	n_{21}	$Np_1 = Nq_1$
$(Z(Nq_1), Z(Nq_2)]$	n_{12}	n_{22}	$Np_2 = N(q_2 - q_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(Z(Nq_s), \infty)$	$n_{1,s+1}$	$n_{2,s+1}$	$Np_{s+1} = N(1 - q_s)$
TOTALS	$n_1 = n$	$n_2 = m$	N

$Z(Nq_i)$ is the $(100q_i)^{\text{th}}$ percentile of the combined sample and, hence, Nq_i of the combined sample points are less than or equal to $Z(Nq_i)$.

Further, if the unknown cpf's of the PS data and the pure-noise data are equal, one expects that n_{1i} is approximately $n_1 p_i = np_i$ and that n_{2i} is approximately $n_2 p_i = mp_i$ for all i.

One then generalizes the contingency-table tests for independence to obtain detector statistics of the form below.

For the sign-quantile detector,

$$Q(n_1, n_2; q_1, \cdots, q_s; r, \Psi)$$

$$= \sum_{i=1}^{s+1} \sum_{j=1}^{2} \left[\left|\frac{n_{ji}}{n_j} - p_i\right| \Psi(n_{ji}/n_j, p_i)\right]^r \qquad (40)$$

where $n_1 = n$, $n_2 = m$ to simplify the summation, and the p_i are such that each Np_i is an integer.

On setting $s = 1$, $q = \frac{1}{2}$, $r = 2$, $\Psi(n_{ji}/n_j , p_i) = \sqrt{n_j/p_i}$ and making the usual chi-square correction for a 2 x 2 contingency table, one obtains the median detector

$$Q_1 = \frac{N}{nm} \left(|2n_{11} - n| - 1 \right)^2 \tag{41}$$

On setting $p_i = (s+1)^{-1}$, $r = 2$ and $\Psi = 1$, one obtains, on simplifying,

$$Q_{s+1} = \frac{(s+1)N}{nm} \sum_{i=1}^{s+1} n_{1i}^2 - \frac{nN}{m} \tag{42}$$

The Matusita modification yields

$$\hat{Q}'_{s+1} = \sum_{i=1}^{s+1} \left(\sqrt{\frac{n_{1i}}{n}} - \sqrt{\frac{n_{2i}}{m}} \right)^2 = 2\left(1 - \frac{1}{\sqrt{nm}} \sum_{i=1}^{s+1} \sqrt{n_{1i}n_{2i}} \right) \tag{43}$$

As an illustration of the computations one might consider the following example.

Example 9.

PFA $\alpha = .005$ $n = 20$; $m = 16$

Q_1 detector: Decide YES iff $\frac{36}{(20)(16)} \left(|2n_{11} - 20| - 1 \right)^2$
$= .450 \left(|n_{11} - 10| - .5 \right)^2 > 7.88$ (the 99.5th percentile of the chi-square distribution with 1 degree of freedom)

Q_4 detector: Decide YES iff $\frac{(4)(36)}{(20)(16)} \sum_{i=1}^{4} n_{1i}^2 - \frac{(20)(36)}{16}$

$= .450 \sum_{i=1}^{4} n_{1i}^2 - 45 > 10.60$ (the 99.5th percentile of the chi-square distribution with 3 degrees of freedom)

$\hat{Q}'_4$ detector: Decide YES iff $2\left(1 - \frac{1}{\sqrt{(20)(16)}} \sum_{i=1}^{4} \sqrt{n_{1i}\, n_{2i}}\right)$

$$= 2 - .112 \sum_{i=1}^{4} \sqrt{n_{1i} n_{2i}} > q(.995)$$

[where $q(.995)$ is the as yet untabulated 99.5th percentile of the distribution of Matusita's[25] two-sample statistic].

PS Data (X_i): -.44, -.10, -.41, 1.08, -.77, -1.65, 1.19, -.55
$n = 20$ -2.05, 1.04, .05, 1.88, .31, -.96, -.18, .52
-1.41, -1.28, .10, -.39

Pure Noise (Y_j): -1.18, -1.31, .62, -.08, -.86, .41, 1.98
$m = 16$ .15, 1.14, -1.95, -.45, 1.28, -1.55, -67
1.18, -.31

r	$Z(r)$	Sample Designation	Computations for Q_1, Q_4, $\hat{Q}'_4$
1	-2.05	X	
2	-1.95	Y	
3	-1.65	X	
4	-1.55	Y	five X's
5	-1.41	X	four Y's
6	-1.31	Y	
7	-1.28	X	
8	-1.18	Y	
9	- .96	X	
10	- .86	Y	
11	- .77	X	
12	- .67	Y	
13	- .55	X	five X's
14	- .45	Y	four Y's
15	- .44	X	
16	- .41	X	
17	- .39	X	
18	- .31	Y	
19	- .18	X	
20	- .10	X	
21	- .08	Y	
22	.05	X	six X's
23	.10	X	three Y's
24	.15	Y	
25	.31	X	
26	.41	Y	
27	.52	X	
28	.62	Y	
29	1.04	X	
30	1.08	X	
31	1.14	Y	four X's
32	1.18	Y	five Y's
33	1.19	X	
34	1.28	Y	
35	1.88	X	
36	1.98	Y	

From observing the ordered values above one can construct the necessary tables for computing Q_1, Q_4, and $\hat{Q}'_4$ (tables 10 and 11).

Table 10

For computing Q_1

Interval \ Sample	X	Y	TOTALS
$(-\infty, Z(18)]$ 1st half	10	8	18
$(Z(18), \infty)$ 2nd half	10	8	18
TOTALS	20	16	36

Table 11

For computing Q_4 and $\hat{Q}'_4$

Interval \ Sample	X	Y	TOTALS
$(-\infty, Z(9)]$ 1st quarter	5	4	9
$(Z(9), Z(18)]$ 2nd quarter	5	4	9
$(Z(18), Z(27)]$ 3rd quarter	6	3	9
$(Z(27), \infty)$ 4th quarter	4	5	9
TOTALS	20	16	36

Using the given formulas one finds:

$Q_1 = .113 < 7.88$ and the Q_1 detector decides NO;

$Q_4 = .90 < 10.60$ and the Q_4 detector decides NO; and

$\hat{Q}'_4 = .022$. (Sinee the $\hat{Q}'_4$ has not been tabulated no deeision can be made here.)

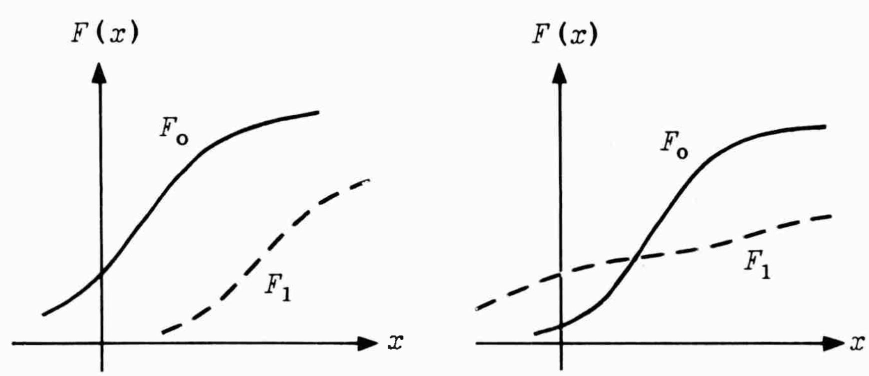
Optimum ehoiees of the number ($s+1$) of intervals and other topies related to goodness criteria will be treated in Section E.

B. MODEL II RUN-BLOCK DETECTORS

This elass of deteetors is also based on the measures of relative "interspersion." An <u>X-run</u> in the ordered eombined sample is a conseeutive sequenee of X's immediately preeeded and immediately succeeded by Y's. A Y-run is similarly defined. If the PS data and the pure-noise data have identieal cpf's, one feels that the two samples should be well interspersed and that, therefore, the total number of runs, i.e., the number of X-runs plus the number of Y-runs, should be relatively large. On the other hand, if the pure-noise (or Y) cpf and the PS (or X) cpf are as illustrated, one expects that the number of runs will be relatively small.

F_0 (PURE NOISE)

F_1 (NOISE PLUS SIGNAL)



A. "Location" alternatives

B. "Spread" alternatives

Once the combined sample is ordered it is a simple matter to calculate the total number of runs. However, it is of some value to give an explicit formula for calculating this total number of runs. To that end, one first notes that

$$\epsilon_{i+1} = \epsilon\left[F_n(Z(i+1)) - F_n(Z(i)) - 1/2n\right] = 1 \text{ iff } Z(i+1) \text{ is an } X;$$

$$\epsilon_{i+1}-\epsilon_i = \begin{cases} 0 \text{ if both } Z(i+1) \text{ and } Z(i) \text{ are } X\text{'s or both are } Y\text{'s} \\ 1 \text{ if } Z(i+1) \text{ is an } X \text{ and } Z(i) \text{ is a } Y \\ -1 \text{ if } Z(i+1) \text{ is a } Y \text{ and } Z(i) \text{ is an } X; \end{cases}$$

$$|\epsilon_{i+1}-\epsilon_i| = \begin{cases} 1 \text{ if } Z(i+1) \text{ and } Z(i) \text{ are from different samples} \\ 0 \text{ if } Z(i+1) \text{ and } Z(i) \text{ are from the same sample;} \end{cases}$$

$$\sum_{i=1}^{n-1} |\epsilon_{i+1}-\epsilon_i| = \text{number of changes of sample } X\text{-to-}Y \text{ and } Y\text{-to-}X \text{ in the ordered combined sample;}$$

and that the total number of runs is therefore

$$\widetilde{R} = 1 + \sum_{i=1}^{n-1} \Big| \epsilon\left[F_n(Z(i+1)) - F_n(Z(i)) - 1/2n\right] - \epsilon\left[F_n(Z(i)) - F_n(Z(i-1)) - 1/2n\right] \Big| \tag{44}$$

Closely related to the run-detector is the detector based on the number of empty Y-blocks. Originally, a Y-block was defined as the interval between two successive Y's in the combined sample. A large number of empty Y-blocks would indicate that the samples were not well interspersed.

The m Y-values $Y_1, \cdots, Y_m$ give rise to $(m+1)$ blocks: $(-\infty, Y(1)]$, $(Y(1), Y(2)]$, $\cdots\cdots$, $(Y(m-1), Y(m)]$, $(Y(m), \infty)$. The number of empty Y-blocks is found to be

$$\mathcal{E} = (m+1) - \sum_{j=1}^{m+1} \epsilon\left[F_n(Y(j)) - F_n(Y(j-1)) - 1/2n\right] \tag{45}$$

where $Y(0) = -\infty$ and $Y(m+1) = +\infty$.

More generally, one may wish to consider blocks determined by certain percentiles of the Y (or pure noise) sample. Let $0 = q_0 < q_1 < \cdots < q_s < q_{s+1} = 1$ be probabilities such that each mq_i is an integer; then a more general empty-block statistic is

$$\mathcal{E}(n, m, q_1, \cdots, q_s) = (s+1)$$

$$- \sum_{j=1}^{s+1} \epsilon \left[F_n\left(Y(mq_j)\right) - F_n\left(Y(mq_{j-1})\right) - 1/2n\right] \qquad (46)$$

The detectors based on statistics (45) and (46) are, of course, immediate extensions of the Model I empty-cell detector. The Model II extensions of the Model I spacing detectors are based on statistics of the form

$$S(n, m; q_1, \cdots, q_s; r, \Psi) = \sum_{j=1}^{s+1} \left| F_n\left(Y[(m+1)q_j]\right)\right.$$

$$\left. - F_n\left(Y[(m+1)q_{j-1}]\right) - (q_j - q_{j-1}) \right|^r \Psi_j \qquad (47)$$

where $0 = q_0 < q_1 < \cdots < q_s < q_{s+1} = 1$.

Corresponding to the Model I situation there are two special cases,

$$S(n, m) = \sum_{j=1}^{m+1} \left| F_n\left(Y(j)\right) - F_n\left(Y(j-1)\right) - \frac{1}{m+1} \right| \qquad (48)$$

$$S_2(n, m) = \sum_{j=1}^{m+1} \left| F_n\left(Y(j)\right) - F_n\left(Y(j-1)\right) - \frac{1}{m+1} \right|^2 \qquad (49)$$

For the final detector discussed in this section, one considers a statistic specifically designed for the one-sided situation just illustrated (A, in preceding figure). It is the Epstein-Rosenbaum

statistic which equals the number of X-values exceeding the last Y-value, i.e.,

$$L' = n\left[1 - F_n\left(Y(m)\right)\right] \tag{50}$$

Since the last order statistic $Y(m)$ is relatively "unstable," one quite often makes use of the more general statistic

$$L(n, m; q) = n\left[1 - F_n\left(Y(mq)\right)\right] \tag{51}$$

the number of X-values exceeding the $100q^{th}$ percentile of the Y-sample.

Before giving an example of the use of the statistics described above it seems worthwhile to comment that except for $\widetilde{R}$, the statistics described above are asymmetric, in the sense that one is working with Y-blocks or Y-runs. Detectors with similar rationales could be constructed if one considered X-blocks and X-runs. Also there is some merit in considering the sum of the statistics based on the X-blocks and the Y-blocks. This latter technique is used in constructing $\widetilde{R}$ which is the sum of the number of X-runs.

Example 10. Consider the data of Example 9 with $n = 20$; $m = 16$; and PFA $\alpha = .005$.

r	$Z(r)$	Cum. No. of Runs	Cum. No. of Empty Y-blocks	L'	$L(n, m, .75)$	$F_n(Y(j))$	$\lvert D_j \rvert = \lvert F_n(Y(j)) - F_n(Y(j-1)) - \frac{1}{17} \rvert$	$10^6 D_j^2$
1	X	1	0					
2	Y	2	0			.05	|.05-0-.059| = .009	81
3	X	3	0					
4	Y	4	0			.10	|.10-.05-.059| = .009	81
5	X	5	0					
6	Y	6	0			.15	|.15-.10-.059| = .009	81
7	X	7	0					
8	Y	8	0			.20	|.20-.15-.059| = .009	81
9	X	9	0					
10	Y	10	0			.25	|.25-.20-.059| = .009	81
11	X	11	0					
12	Y	12	0			.30	|.30-.25-.059| = .009	81
13	X	13	0					
14	Y	14	0			.35	|.35-.30-.059| = .009	81
15	X	15	0					
16	X	15	0					
17	X	15	0					
18	Y	16	0			.50	|.50-.35-.059| = .091	8281
19	X	17	0					
20	X	17	0					
21	Y	18	0			.60	|.60-.50-.059| = .041	1681
22	X	19	0					
23	X	19	0					
24	Y	20	0			.70	|.70-.60-.059| = .041	1681
25	X	21	0					
26	Y	22	0			.75	|.75-.70-.059| = .009	81
27	X	23	0					
28	Y	24	0			.80	|.80-.75-.059| = .009	81
29	X	25	0		1			
30	X	25	0		2			
31	Y	26	0		2	.90	|.90-.80-.059| = .041	1681
32	Y	26	1		2	.90	|.90-.90-.059| = .059	3481
33	X	27	1		3			
34	Y	28	1		3	.95	|.95-.90-.059| = .009	81
35	X	29	1		4			
36	Y	30	1	0	4	1.00	|1.00-.95-.059| = .009	81
							.372	17,696
		$\tilde{R} = 30$	$\varepsilon = 1$	$L' = 0$	$L(n, m, .75) = 4$		$S(n, m) = .372$	$S_2(n, m) = .018$

The decision rules are:

$\tilde{R}$ detector: Decide YES iff $\tilde{R} \leq 10$.

$\mathcal{E}$ detector: Decide YES iff $\dfrac{\mathcal{E} - \mu(\mathcal{E})}{\sigma(\mathcal{E})} > +2.576$

where $\mu(\mathcal{E}) = n^2/N = 11.11$

$\sigma(\mathcal{E}) = nm/N^{3/2} = 1.48$

L' detector: Decide YES iff $20\,[1 - F_n\,(Y(16))] \geq 7$.
(Note: The value "7" corresponds to PFA α = .01. There are no tabulated values for α = .005 as of this writing.)

$L(n, m, .75)$ detector: Decide YES iff $20\,[1 - F_n\,(Y(12))] \geq n[1 - p(\alpha)] + 1$ where $[q - q(\alpha)] \cdot \left\{\dfrac{p(\alpha)\,[1 - p(\alpha)]}{n} + \dfrac{q(1 - q)}{m}\right\}^{-\frac{1}{2}} = \Phi^{-1}(\alpha)$ and, consequently, $n\,[1 - p(\alpha)] + 1 = 15$.

$S(n, m)$ detector: Decide YES iff $\displaystyle\sum_{j=1}^{17} \left| F_n\,(Y(j)) - F_n\,(Y(j\text{-}1)) - \frac{1}{17} \right| \geq b_1$, where b_1 is the untabulated 99.5th percentile of the $S(n, m)$ distribution.

$S_2(n, m)$ detector: Decide YES iff $\displaystyle\sum_{j=1}^{17} \left| F_n\,(Y(j)) - F_n\,(Y(j\text{-}1)) - \frac{1}{17} \right|^2 \geq c_1$, where c_1 is the untabulated 99.5th percentile of the $S_2(n, m)$ distribution.

Consequently, the $\tilde{R}$, $\mathcal{E}$, L', and $L(n, m, .75)$ detectors decide YES. Further, since b_1 and c_1 are unknown, the decisions of the $S(n, m)$ and $S_2(n, m)$ cannot be determined.

Optimum choices of q and other topics related to goodness criteria will be treated in Section III E.

C. MODEL II RANK-SUM DETECTORS

The detectors of this type are generalizations of the detector based on the statistic $V = \sum_{i=1}^{n} R(X_i)$. The intuitive justification for the use of this statistic is as follows. If $F_0 > F_1$ (see illustration, p. 41), then the X distribution is stochastically larger than the Y distribution, and one expects that in the ranked combined sample the X values will occupy the upper ranks. Hence $V = \frac{1}{n} \sum_{i=1}^{n} R(X_i)$ will be significantly large.

In order to determine which values of $\sum R(X_i)$ are "significantly large" or "significantly small," one needs tables other than those previously used. Therefore, one might ask which constants can be used in place of the ranks, to permit use of the "standard tables" and satisfy certain goodness criteria. One, then, is going to consider statistics of the form

$$V_h = \frac{1}{n} \sum_{i=1}^{n} h\,[R(X_i)] \tag{52}$$

where h is an appropriately chosen function. Table 12 presents some usual choices of h.

Table 12. Statistics $\frac{1}{n} \sum_{i=1}^{n} h\,[R(X_i)]$

Corres. 1-Sample Statistic	Two-Sample Statistic	$h_1(r)$ (Expectations)	$h_2(r)$ (Percentiles)
$\bar{Z}$	Fisher-Yates-Van der Waerden	$E\,[\xi(r) \mid \Phi]$	$\Phi^{-1}\left(\frac{r}{N+1}\right)$
$\bar{U}$	Mann-Whitney-Wilcoxon	$E\,[\xi(r) \mid U_o] = \frac{r}{N+1}$	$U_o^{-1}\left(\frac{r}{N+1}\right) = \frac{r}{N+1}$
π'	Savage	$E[\xi(r) \mid H_e] = \sum_{j=N-r+1}^{N} 1/j$	$H_e^{-1}\left(\frac{r}{N+1}\right) = \ln\left[\frac{N+1}{N+1-r}\right]$

Where Φ, U_o and H_e are, respectively, the cpf's of the standard normal, standard uniform, and standard exponential distributions (see Section II C); $\xi(r)$ is the rth order statistic of a random sample $\{W_1, W_2, \cdots, W_n\}$ from a population with cpf H; and $E[W(r) \mid H]$ is its expectation.

Historically, the Fisher-Yates detector was studied first. In an attempt to attain approximate or asymptotic normality one replaces each $R(X_i)$ by $E\{W[R(X_i)] \mid \Phi\}$ to obtain the statistic

$$V'(\Phi) = \frac{1}{n} \sum_{i=1}^{n} E\{W[R(X_i)] \mid \Phi\}. \tag{53}$$

This statistic turns out to be asymptotically normal, but for small or moderate values of n and m, special tables of significance levels are needed. Further, one needs the not-too-common tables of expected values $E[W(r) \mid \Phi]$ of normal order statistics. To circumvent this latter difficulty (which does not exist for U_o and H_e since explicit formulas are available in these cases), one introduces the statistic based on percentiles $\Phi^{-1}\left(\frac{r}{N+1}\right)$, i.e.,

$$V(\Phi) = \frac{1}{n} \sum_{i=1}^{n} \Phi^{-1}\left(\frac{R(X_i)}{N+1}\right). \tag{54}$$

Since there is no apparent reason for restricting consideration to statistics related to Φ, one naturally extends consideration to statistics based on U_o and H_e and, finally, to those based on arbitrary strictly increasing cpf's. The general rank-sum detectors are, hence, based on statistics of the form

$$V'(H) = \frac{1}{n} \sum_{i=1}^{n} E\{W[R(X_i)] \mid H\} \text{ and} \tag{55}$$

$$V(H) = \frac{1}{n} \sum_{i=1}^{n} H^{-1}\left[\frac{R(X_i)}{N+1}\right]. \tag{56}$$

Before illustrating the necessary calculations, one should note the following:

1. In practice, one usually employs principally those detectors for which $H = \Phi$, U_o or H_e. These restrictions, as in the one-sample case, coincide with certain seemingly unrelated information theory results.

2. The statistics $V(H)$ and $V'(H)$ are primarily designed for the one-sided situations as illustrated on page 41. Several detectors for other types of situations will be discussed after the numerical example.

3. In the one-sample case, the corresponding rank-sum detectors are based on statistics of the form

$$K(H) = \frac{1}{n} \sum_{i=1}^{n} H^{-1}[F_o(X_i)] .$$

Since $F_o(X_i)$ is the percentile rank of X_i, one might be led to consider statistics of the form

$$\frac{1}{n} \sum_{i=1}^{n} H^{-1}[R(X_i)] .$$

This statistic is not well defined, since H^{-1} is only defined for numbers between 0 and 1. The next step in the evolution is the detector based on

$$\frac{1}{n} \sum_{i=1}^{n} H^{-1}\left[\frac{R(X_i)}{N}\right],$$

which would be well defined except for the fact that $H^{-1}(1) = \infty$, if H is a strictly increasing continuous cpf. The final result is, then, the given statistic $V(H)$.

4. For the case $H = U_o$, $U_o^{-1}(u) = u$ whenever $0 < u < 1$ and, hence,

$$V(U_o) = \frac{1}{n} \sum_{i=1}^{n} \frac{R(X_i)}{N+1} = \frac{1}{n(N+1)} \sum_{i=1}^{n} R(X_i) \qquad (57)$$

which is equivalent to the more usual

$$\frac{1}{n} \sum_{i=1}^{n} R(X_i) \quad \text{or} \quad \sum_{i=1}^{n} R(X_i)\ ,$$

the Mann-Whitney-Wilcoxon statistic. Further, $V(U_0) = V'(U_0)$.

Finally, to facilitate the computations in the example below, one needs to know the following theorem (see ref. 8, p. 500):

Theorem 7. If $F_0 = F_1$, i.e., if the PS data are pure noise, then

(i) the statistic $V_h = \frac{1}{n} \sum_{i=1}^{n} h\,[R(X_i)]$ has

$$\text{mean } E(V_h) = \frac{1}{N} \sum_{r=1}^{N} h(r)$$

and variance

$$\text{var}\,(V_h) = \frac{m}{nN^2}\left\{\left(\frac{N}{N-1}\right) \sum_{r=1}^{N} h^2(r) - \frac{1}{N-1}\left[\sum_{r=1}^{N} h(r)\right]^2\right\}$$

and

(ii) the statistics

$$\left\{\frac{V(H) - E\,[V(H)]}{\sqrt{\text{var}\ [V(H)]}}\right\} \text{ and } \left\{\frac{V'(H) - E\,[V'(H)]}{\sqrt{\text{var}\ [V(H)]}}\right\}$$

have asymptotically a normal distribution for $H = \Phi$, U_0, and H_e.

Example 11. Consider the situation for which $n = 10$, $m = 8$, $N = 18$, PFA $\alpha = .01$ and the data are as given.

PS data (X_i): 18.1, 14.3, 15.2, 14.5, 17.9, 17.5, 15.3, 16.5, 17.1, 16.6

Pure-noise data (Y_j): 14.2, 15.4, 16.4, 17.7, 18.2, 17.3, 15.1, 16.1

Using normal approximations, one derives the decision rules below for the detectors described above and for the one-sided alternatives illustrated on page 41.

$V(U_o)$ detector:

Decide YES iff $\dfrac{V(U_o) - 9.5}{1.407} < -2.326.$

$V(\Phi)$ detector:

Decide YES iff $\dfrac{V(\Phi) - 0}{.232} < -2.326.$

$V(H_e)$ detector:

Decide YES iff $\dfrac{V(H_e) - 0.873}{.254} < -2.326.$

$V'(H_e)$ detector:

Decide YES iff $\dfrac{V'(H_e) - 1}{.169} < -2.326.$

$V'(\Phi)$ detector:

Decide YES iff $\dfrac{V'(\Phi) - 0}{.243} < -2.326.$

The ranked data and computations are best given in tabular form.

r	$Z(r)$	Sample	(Mann-Whitney -Wilcoxon) $(N+1)\ V(U_o)$ $h_1(r) = r$	(Van de Waerden) $V(\Phi)$ $h_2(r) = \Phi^{-1}\left(\frac{r}{N+1}\right)$	(Fisher-Yates) $V'(\Phi)$ $h_3(r) = E\left[\xi(r) \mid \Phi\right]$	$V(H_e)$ $h_4(r) = \ln(N-1) - \ln(N+1-r)$	$V'(H_e)$ $h_5(r) = \sum_{j=N-r+1}^{N} 1/j$
1	14.2	Y	1	-1.62	-1.82	.054	.056
2	14.3	X	2	-1.25	-1.35	.111	.115
3	14.5	X	3	-1.00	-1.07	.171	.177
4	15.1	Y	4	-0.80	-0.85	.236	.244
5	15.2	X	5	-0.63	-0.67	.305	.315
6	15.3	X	6	-0.48	-0.50	.379	.392
7	15.4	Y	7	-0.34	-0.35	.459	.476
8	16.1	Y	8	-0.20	-0.21	.546	.566
9	16.4	Y	9	-0.07	-0.07	.641	.666
10	16.5	X	10	0.07	0.07	.747	.777
11	16.6	X	11	0.20	0.21	.864	.902
12	17.1	X	12	0.34	0.35	.998	1.040
13	17.3	Y	13	0.48	0.50	1.152	1.212
14	17.5	X	14	0.63	0.61	1.335	1.412
15	17.7	Y	15	0.80	0.85	1.558	1.662
16	17.9	X	16	1.00	1.07	1.845	1.995
17	18.1	X	17	1.25	1.35	2.251	2.495
18	18.2	Y	18	1.62	1.82	2.944	3.495
$\Sigma h(r)$			171	0	0	16.60	18.00
$\Sigma h^2(r)$			2109	13.23	15.75	26.23	32.50
$E(V_h)$			9.5	0	0	.922	1.00
$\text{var}(V_h)$			1.979	.0540	.0643	.0286	.0592
$\sqrt{\text{var}(V_h)}$			1.407	.232	.254	.169	.243
V_h			9.6	.013	.013	.901	.962
$\frac{V_h - E(V_h)}{\sqrt{\text{var}(V_h)}}$			.071	.056	.051	-.124	-.156

Therefore, each of the detectors decides NO.

One should note here the following useful lemma.

Lemma. If U is a random variable with cpf U_o, then for each strictly increasing continuous cpf H

$H^{-1}(U)$ is a random variable with cpf H; and

$E\{g[H^{-1}(U)] \mid U_o\} = E[g(w) \mid H]$ whenever such expectations exist.

Consequently, one can rewrite the statistics

$$V'(H) = \frac{1}{n} \sum_{i=1}^{n} E\{W[R(X_i)] \mid H\}$$

as

$$V'(H) = \frac{1}{n} \sum_{i=1}^{n} E(H^{-1}\{U[R(X_i)] \mid U_o\} . \tag{58}$$

Recent results[13,14] indicate that the statistics $V'(H)$ are special cases of the more general class of "locally most powerful" rank statistics. They are of the form

$$V(h,\theta) = \frac{1}{n} \sum_{i=1}^{n} E\left(J_o'\{U[R(X_i)]\}\right) \tag{59}$$

where

$U(i)$ is the i^{th} order statistic of a sample of size N from U_o;

$$J_o'(u) = \frac{\partial}{\partial\theta} \ln h_\theta'(u)\Big|_{\theta=\theta_o}$$

$\{h_\theta\}$ is a family of cpf's on the unit interval (0,1) such that

$h_{\theta_o}(u) \equiv u$, and $h_\theta{}' = \dfrac{d h_\theta}{du}$ is non-zero for all u and θ.

[Model II detectors based on these statistics correspond to the Model I detectors discussed immediately preceding Example 7 (p. 45)].

Special examples of these statistics are

$$\frac{1}{n} \sum_{i=1}^{n} E\left[\left(\Phi^{-1}\{U[R(X_i)]\}\right)^2 \,\middle|\, U_o\right] \tag{60}$$

$$\frac{1}{n} \sum_{i=1}^{n} E\left[\frac{\left(H_C{}^{-1}\{U[R(X_i)]\}\right)^2}{1+\left(H_C{}^{-1}\{U[R(X_i)]\}\right)^2} \,\middle|\, U_o\right] \tag{61}$$

and

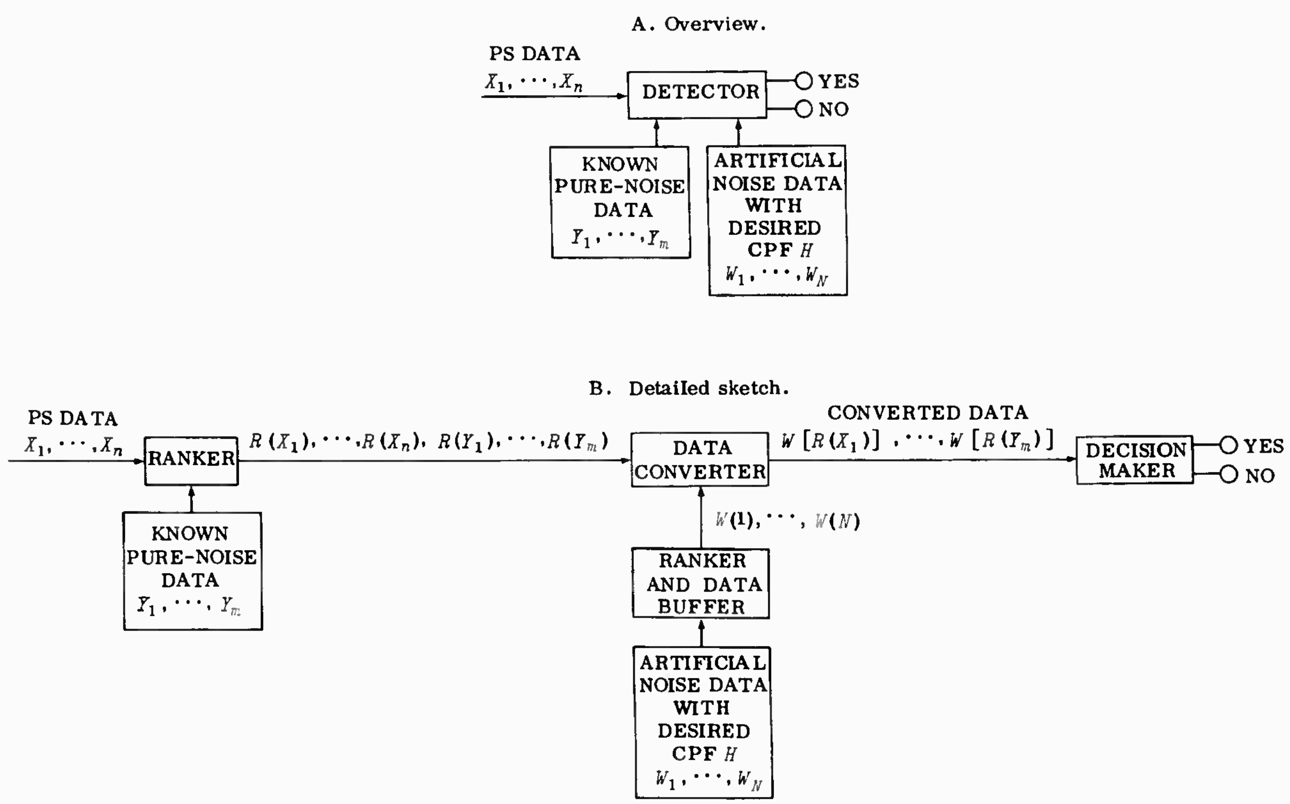
$$\frac{1}{n} \sum_{i=1}^{n} E\left[\frac{H_C{}^{-1}\{U[R(X_i)]\}}{1+\left(H^{-1}\{U[R(X_i)]\}\right)^2} \,\middle|\, U_o\right] \tag{62}$$

H_C is the cpf with derivative $[\pi(1+x^2)]^{-1}$.

The figure on page 76 illustrates (A) detectors based on equation 62, primarily designed for "location" alternatives, and (B) detectors based on equations 60 and 61, used principally for "spread" alternatives. In each of these cases there are difficulties in computing and tabulating the expectations $E[W(r) \mid H]$, as well as in attaining predetermined PFA α's. To prevent these difficulties, one should consider the detectors which make use of artificial noise generators, as discussed and diagrammed in the following section.

D. MODEL II ARTIFICIAL NOISE DETECTORS

In attempting to obtain the desired preassigned PFA and to tabulate the distributions involved, one should consider an adaptation of a statistical result reported by Doksum.[11]



Detector employing artificial noise.

The mathematical basis of the method to be described is theorem 8, which is proved in reference 11.

Theorem 8. If $Z_1, \cdots, Z_N$ and $W_1, \cdots, W_N$ are independent random samples with common continuous cpf's G and H, respectively, then $W[R(Z_1)], \cdots, W[R(Z_N)]$ (i) has the statistical distribution of a random sample with common cpf H, and
(ii) $R\{W[R(Z_i)]\} = R(Z_i)$ for $i = 1, 2, \cdots, N$.

The result of theorem 8 is that on replacing each Z-value with the W-value of the same rank, one is able to replace a random sample with cpf G by a random sample with cpf H, while preserving the ranks on which the Model II detectors will base their YES-NO decisions.

Since the Model II detectors under consideration are rank detectors, and since the described transformation preserves the ranks, there is no loss of information, but a gain of tractability, convenience, and economy in using the transformation with a judicious choice of the W cpf H.

As just illustrated,

1. the combined sample $X_1, \cdots, X_n$; $Y_1, \cdots, Y_m$ is ranked;

2. a random sample $W_1, W_2, \cdots, W_N$ of artificial noise with common cpf H is generated;

3. each combined sample value X_i or Y_j is replaced by the W-value $W[R(X_i)]$ or $W[R(Y_j)]$ which has the same rank;

4. an appropriate statistic is evaluated with the W inputs $W(X_1), \cdots, W(X_n)$; $W(Y_1), \cdots, W(Y_m)$; and

5. on the basis of this computed value, the detector decides YES or NO.

(To make use of such a detection procedure, one needs, in addition to the equipment used by previously described detector models, a random noise generator capable of generating a W sample having the desired cpf H.)

The major question at this point is: Which statistics are appropriate for the transformed data?

To answer this question one first considers statistics of the form

$$V'(H) = \frac{1}{n} \sum_{i=1}^{n} E\left\{W[R(X_i)] \mid H\right\}$$

Since for fixed n, m, and hence for N, the N expected values $E[W(1) \mid H], \cdots, E[W(N) \mid H]$ are fixed, the set of values $E\left\{W[R(X_1)] \mid H\right\}, \cdots, E\left\{W[R(X_n)] \mid H\right\}$ uniquely determines and is determined by the $E\left\{W[R(Y_1)] \mid H\right\}, \cdots, E\left\{W[R(Y_m)] \mid H\right\}$. Therefore, the statistic $V'(H)$ is equivalent to the statistic

$$V''(H) = \frac{1}{n} \sum_{i=1}^{n} E\left\{W[R(X_i)] \mid H\right\} - \frac{1}{m} \sum_{j=1}^{m} E\left\{W[R(Y_j)] \mid H\right\}$$

For non-artificial noise detectors it is more convenient to work with $V'(H)$ than with $V''(H)$, since the former statistic requires less calculation and is equivalent to the latter.

For artificial noise detectors, it is proposed that the expectation signs "E" be dropped and that one consider statistics of the form

$$\widetilde{T}(H) = \frac{1}{n} \sum_{i=1}^{n} W[R(X_i)] \tag{63}$$

and

$$T(H) = \frac{1}{n} \sum_{i=1}^{n} W[R(X_i)] - \frac{1}{m} \sum_{j=1}^{m} W[R(Y_j)] \tag{64}$$

where $W_1, \cdots, W_N$ is a random sample of artificial noise generated with common cpf H. Since the values $W(1), \cdots, W(N)$ are not fixed as are the $E[W(1) \mid H], \cdots, E[W(N) \mid H]$, the statistics $\widetilde{T}(H)$ and $T(H)$ are not equivalent as are $V'(H)$ and $V''(H)$. It is clear that

$\widetilde{T}(H)$ does not make use of all the randomly transformed data; and it has been proved[11] that under some usual (translation) alternatives and for large sample sizes, detectors based on $\widetilde{T}(H)$ have approximately the fraction $\frac{n}{N}$ of the PFD β of detectors based on $T(H)$. For that reason attention will be restricted to detectors based on the statistics $T(H)$.

Before illustrating the technique outlined above it is necessary to state the following theorem, which will be employed in the example.

Theorem 9. If the pure-noise and PS-data cpf's are equal, and $W_1, \cdots, W_N$ is a random sample generated by the artificial noise generator with cpf H, then $T(H)$ has the distribution of the difference of means of independent random samples of size n and m, respectively from populations with cpf's equal to H. In particular,

(i) if $H = \Phi, \sqrt{nm/N}\ T(\Phi)$ has cpf Φ;

(ii) if $H = U_o$, and $n = m = N/2$, then $.5\,[1 + T(U_o)]$ has the cpf of the mean of a random sample of size N from U_o, i.e., its density is

$$N^N/(N-1)! \sum (-1)^r \binom{N}{r} \left(x - \frac{r}{N}\right)^{N-1}$$

where the summation is over $r \leq Nx$;

(iii) if $H = H_e$, and $n = m = N/2$, then $N\ T(H_e)$ has the cpf of the difference of two independent chi-square variables, each with N degrees of freedom; and

(iv) $\sqrt{3N}\ T(U_o)$ and $\sqrt{n/2}\ T(H_e)$ are asymptotically $N(0,1)$ when $n = m = N/2$.

It should be remarked that with reference to (iv) of Theorem 9, the normal approximation is "adequate" for $n = m = N/2 \geq 10$.

Example 12. Consider the situation for which $n = m = 10$, $N = 20$, PFA $\alpha = .05$ and the data are as given.

PS data (X_i): 18.1, 14.3, 15.2, 14.5, 17.9, 15.3, 16.5, 17.1, 16.6, 18.4

Pure-noise data (Y_j): 18.3, 17.5, 14.2, 15.4, 16.4, 17.7, 18.2, 17.3, 15.1, 16.1

Using results (i) and (iv) of theorem 9, one derives the following decision rules for the case ($F_o > F_1$) illustrated on page 41.

$T(\Phi)$ detector: Decide YES iff $\sqrt{nm/N}\ T(\Phi) = 2.236\ T(\Phi) < -1.645$

$T(U_o)$ detector: Decide YES iff $\sqrt{3N}\ T(U_o) = 7.746\ T(U_o) < -1.645$

$T(H_e)$ detector: Decide YES iff $\sqrt{n/2}\ T(H_e) = 2.236\ T(H_e) < -1.645$

The detectors above need artificial noise samples generated from Φ, U_o and H_e distributions, respectively. One such collection of samples is as follows.

<u>Φ artificial noise sample</u>: 2.455, -0.531, -0.634, 1.279, 0.046, -0.525, 0.007, -0.162, -1.618, 0.378, -0.057, 1.356, -0.918, 0.012, -0.911, 1.237, -1.384, -0.959, 0.731, 0.717

<u>U_o artificial noise sample</u>: .950, .455, .317, .869, .358, .853, .540, .985, .266, .373, .920, .164, .998, .073, .495, .496, .641, .417, .906, .903

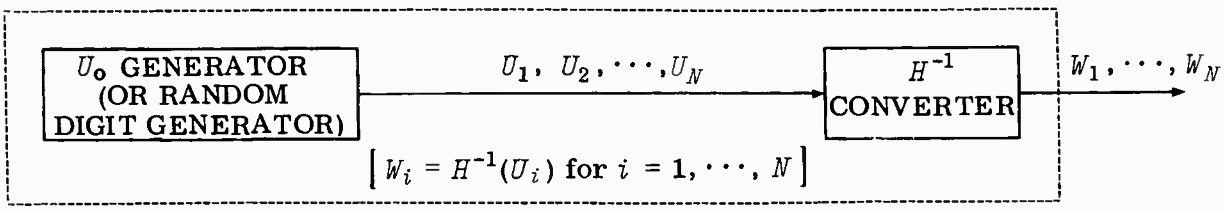
<u>H_e artificial noise sample</u>: .623, .272, .511, .294, .257, .639, .818, .165, .110, .691, .606, .317, .219, .428, .223, .275, .051, .544, .492, .466

Table 13

r	$Z(r)$	Sample	Artificial Noise $\{W_i\}$ for $T(\Phi)$	Artificial Noise $\{W_i\}$ for $T(U_o)$	Artificial Noise $\{W_i\}$ for $T(H_e)$
1	Y	14.2	-1.618	.073	.051
2	X	14.3	-1.384	.164	.110
3	X	14.5	-0.959	.266	.165
4	Y	15.1	-0.918	.317	.219
5	X	15.2	-0.911	.358	.223
6	X	15.3	-0.634	.373	.257
7	Y	15.4	-0.531	.417	.272
8	Y	16.1	-0.525	.455	.275
9	Y	16.4	-0.162	.495	.294
10	X	16.5	-0.057	.496	.317
11	X	16.6	0.007	.540	.428
12	X	17.1	0.012	.641	.466
13	Y	17.3	0.046	.853	.492
14	Y	17.5	0.378	.869	.511
15	Y	17.7	0.717	.903	.544
16	X	17.9	0.731	.906	.606
17	X	18.1	1.237	.920	.623
18	Y	18.2	1.279	.950	.639
19	Y	18.3	1.356	.985	.691
20	X	18.4	2.455	.998	.818
$\sum_{i=1}^{N} W_i$			.519	11.979	8.001
$\sum_{1}^{n} W[R(X_i)]$			.497	5.662	4.013
$\sum_{1}^{m} W[R(Y_j)]$			.022	6.317	3.988
$T(H)$			.0475	-.0655	.0025

On the basis of these computations, all three of these detectors decide NO.

For generating artificial noise with a given strictly increasing continuous cpf H, it is often most convenient to make use of the lemma on page 88. One first generates uniform random numbers, $U_1, \cdots, U_N$, i.e., random numbers with cpf U_0; and then transforms them using H^{-1} to obtain the sample $H^{-1}(U_1), \cdots, H^{-1}(U_N)$ which is statistically equivalent to artificial noise from an "H-generator."



Artificial H–noise generator.

From these considerations one sees that the statistics

$$T(H) = \frac{1}{n} \sum_{1}^{n} W[R(X_i)] - \frac{1}{m} \sum_{1}^{m} W[R(Y_j)]$$

can be rewritten as

$$T(H) = \frac{1}{n} \sum_{1}^{n} H^{-1}\{U[R(X_i)]\} - \frac{1}{m} \sum_{1}^{m} H^{-1}\{U[R(Y_j)]\} \tag{65}$$

Corresponding to the results mentioned in the latter parts of sections II B and III C, it is found[12] that detectors based on the statistics $T(H)$ are special cases of detectors based on statistics of the following forms:

$$T(h, \theta) = \frac{1}{n} \sum_{i=1}^{n} J_\theta\{U[R(X_i)]\} - \frac{1}{m} \sum_{j=1}^{m} J_\theta\{U[R(Y_j)]\}\ ; \tag{66}$$

and

$$T'(h,\theta_0) = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial}{\partial\theta} J_\theta \{U[R(X_i)]\} \Big|_{\theta=\theta_0}$$

$$- \frac{1}{m} \sum_{j=1}^{m} \frac{\partial}{\partial\theta} J_\theta \{U[R(Y_j)]\} \Big|_{\theta=\theta_0} \qquad (67)$$

where

$J_\theta(u) = \ln\ h_\theta'(u)$;

$U(1) < \cdots < U(N)$ are the order statistics of a random sample from U_0; and

$\{h_\theta\}$ is a family of cpf's on the unit interval such that $h_\theta' = \frac{d h_\theta}{du}$ exists and is positive on the unit interval, and $h_{\theta_0}(u) = u$.

The most common versions of these statistics are with

$$\frac{\partial}{\partial\theta} J_\theta(u) \Big|_{\theta=\theta_0} = \frac{F''(F^{-1}(u))}{F'(F^{-1}(u))} ,$$

and with

$$\frac{\partial}{\partial\theta} J_\theta(u) \Big|_{\theta=\theta_0} = \frac{[F''(F^{-1}(u))]\ F^{-1}(u)}{F'(F^{-1}(u))}$$

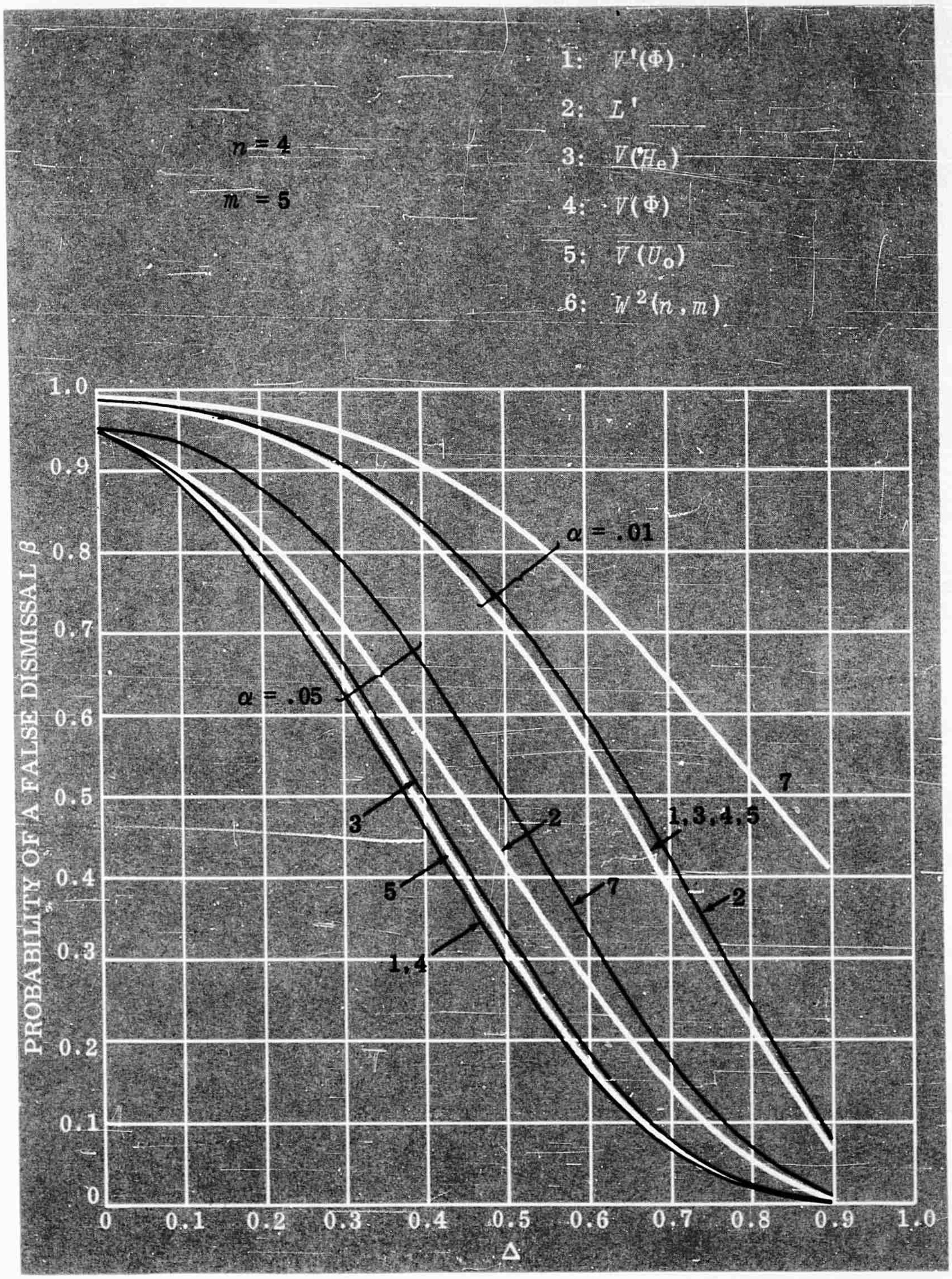
where F is an arbitrary strictly increasing cpf with two derivatives. The first of these two types is primarily for "location" alternatives, and the detectors based on statistics of the second type are for "spread" alternatives.

E. GOODNESS CRITERIA FOR MODEL II DETECTORS

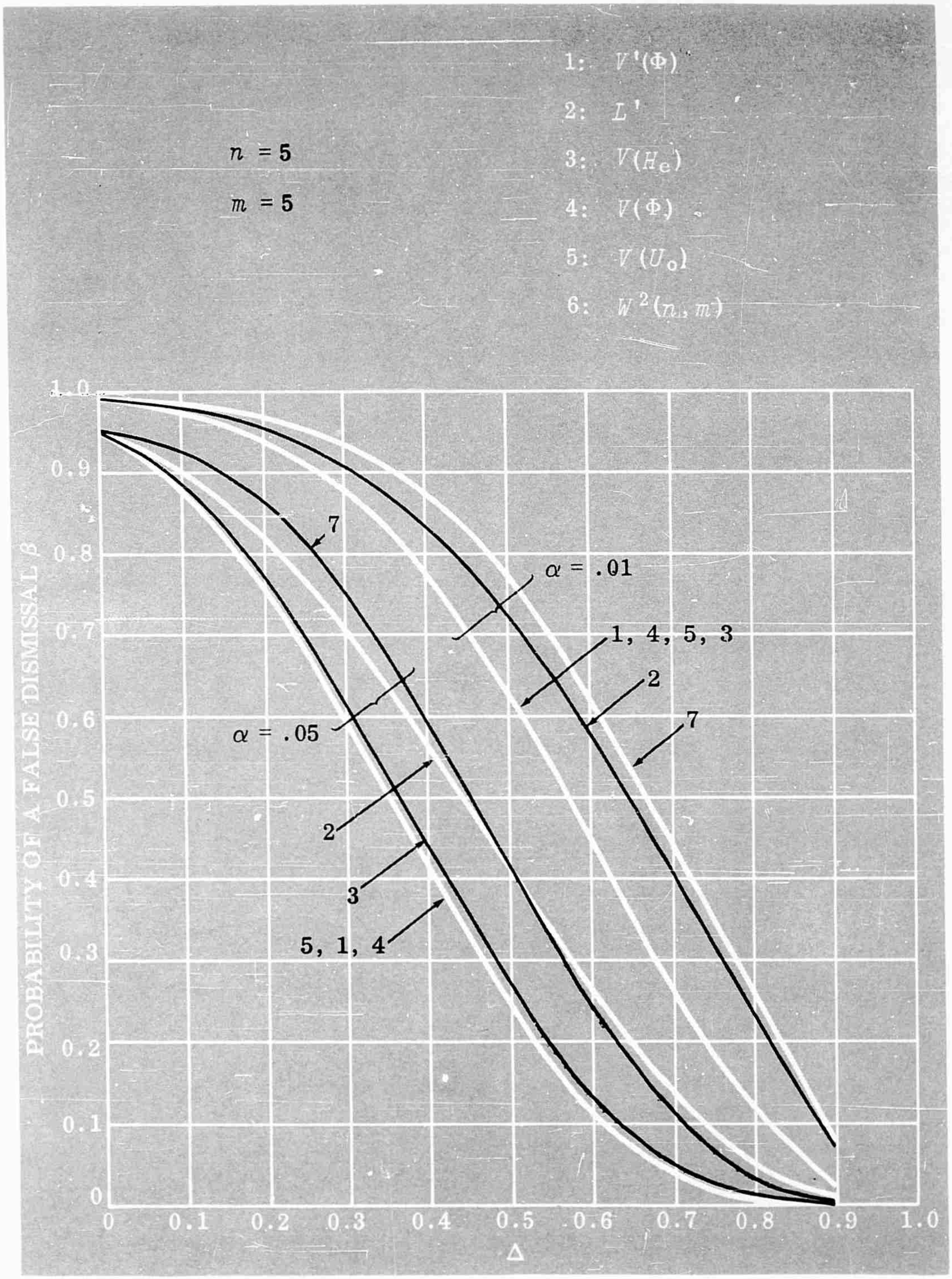
As was the case for Model I detectors, all goodness criteria will be based on PFD β (for fixed PFA α), and since these Model II detectors are SDF, the PFD will be a function solely of $F_0 F_1^{-1}$. Further, because of the multitude of possible alternatives it is physically impossible to make a "reasonably" complete tabulation of PFD. Consequently, one is led to a comparison of detectors based on (1) max-min power for one-sided bands; (2) PFD for Rayleigh alternatives; (3) ARE; and (4) LM PFD (locally minimum PFD) and ALM PFD (asymptotically LM PFD).

1. Minimum PFD for One-sided Bands

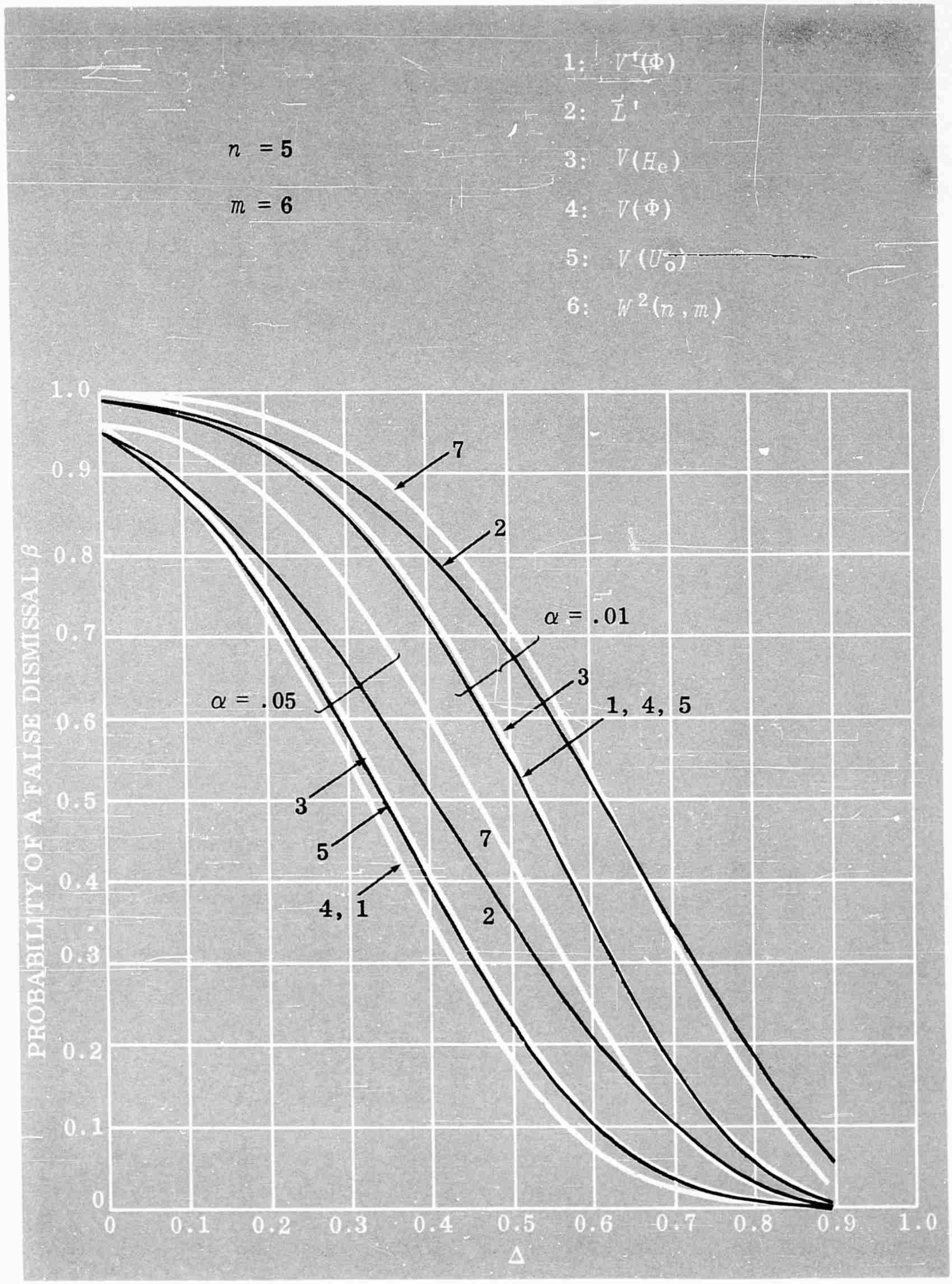
The results here are adaptations of the statistical work of Bell, Moser, and Thompson,[10] which is a partial extension of the statistical work of Chapman.[9] One first considers, for each strictly increasing continuous pure-noise cpf F_0, the family $\mathcal{G}(F_0, \Delta)$ of noise-plus-signal cpf's F_1 satisfying (1) $F_1 < F_0$ and (2) $\max_x [F_0(x) - F_1(x)] = \Delta$. (An illustration is given on p. 47.) In order to apply the max-min PFD criteria one must compute both the maximum and the minimum PFD for F_0 vs. noise-plus-signal cpf's in the one-sided-band family $\mathcal{G}(F_0, \Delta)$. Unfortunately, Bell, Moser, and Thompson[10] were unable to derive a reasonable computation procedure for maximum PFD. The only results available are those for minimum PFD, as given in the graphs on pages 99-111. Here the PFA = .01 or .05; β is plotted vs. Δ; and the sample sizes range from $n = 4$, $m = 5$ to $n = 10$, $m = 10$. These graphs are based exactly on the complete set of computations of Bell, Moser, and Thompson, with the same cases omitted. Those labeled A through L represent the results of the use of accurate approximation formulas, while the curves in the M graph are the result of Monte Carlo calculations for $T(\Phi)$ with 1000 trials per point. It is felt that since other Monte Carlo computations in this study are based on 10,000 trials per point, the results for $T(\Phi)$ are somewhat less accurate than the other Monte Carlo computations.



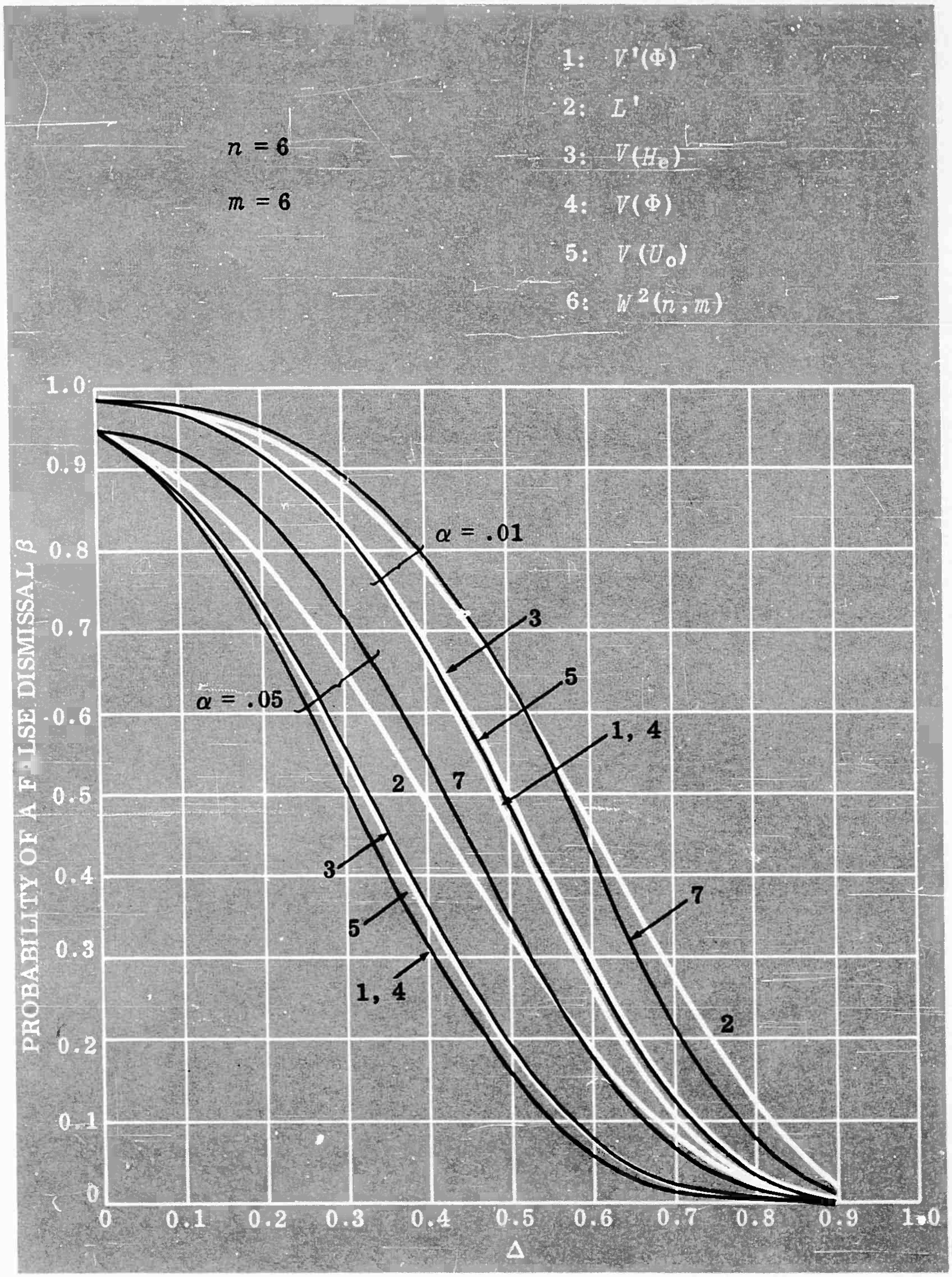
Minimum PFD for one-sided noise.



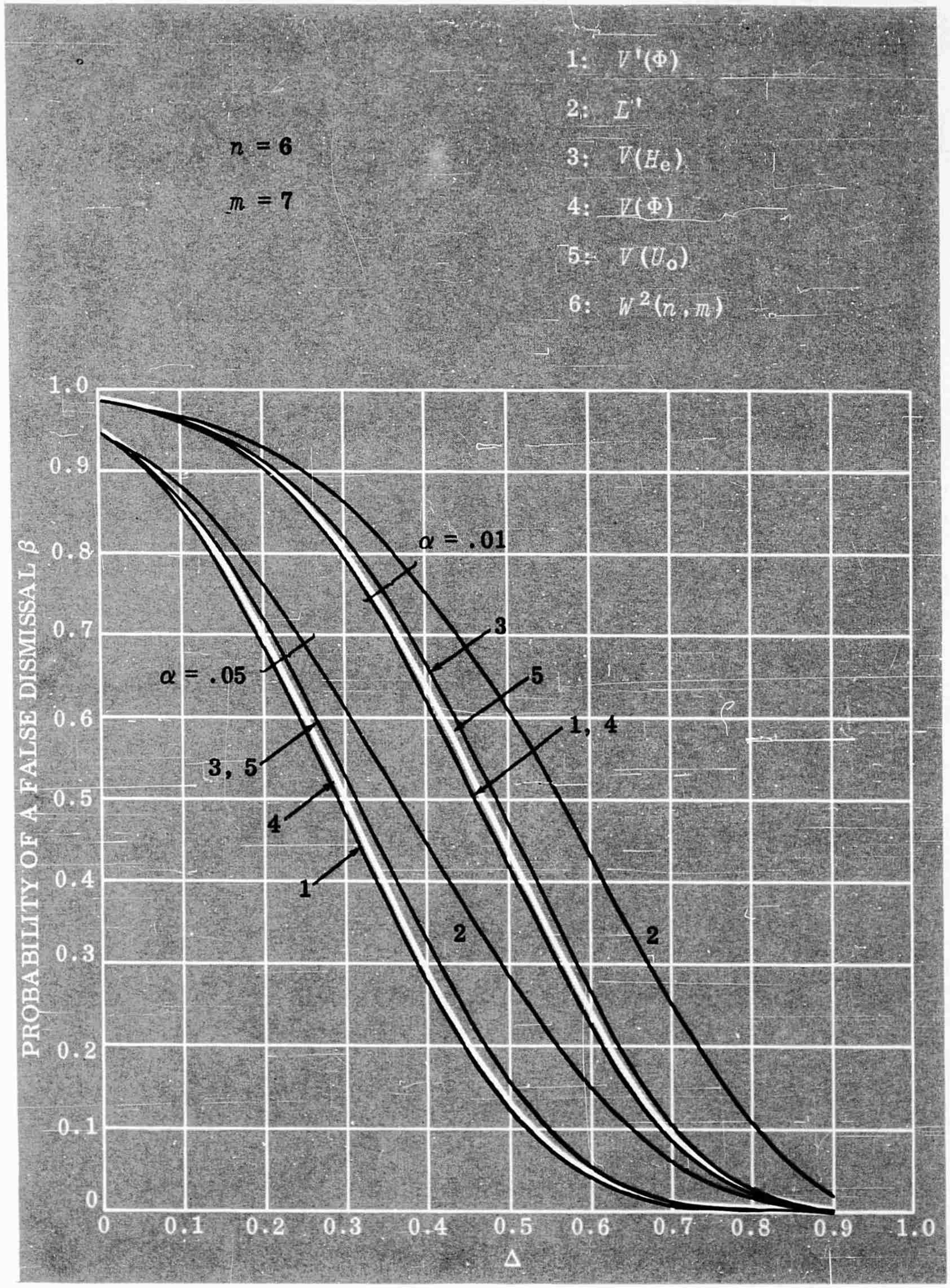
Minimum PFD for one-sided noise. (Continued)



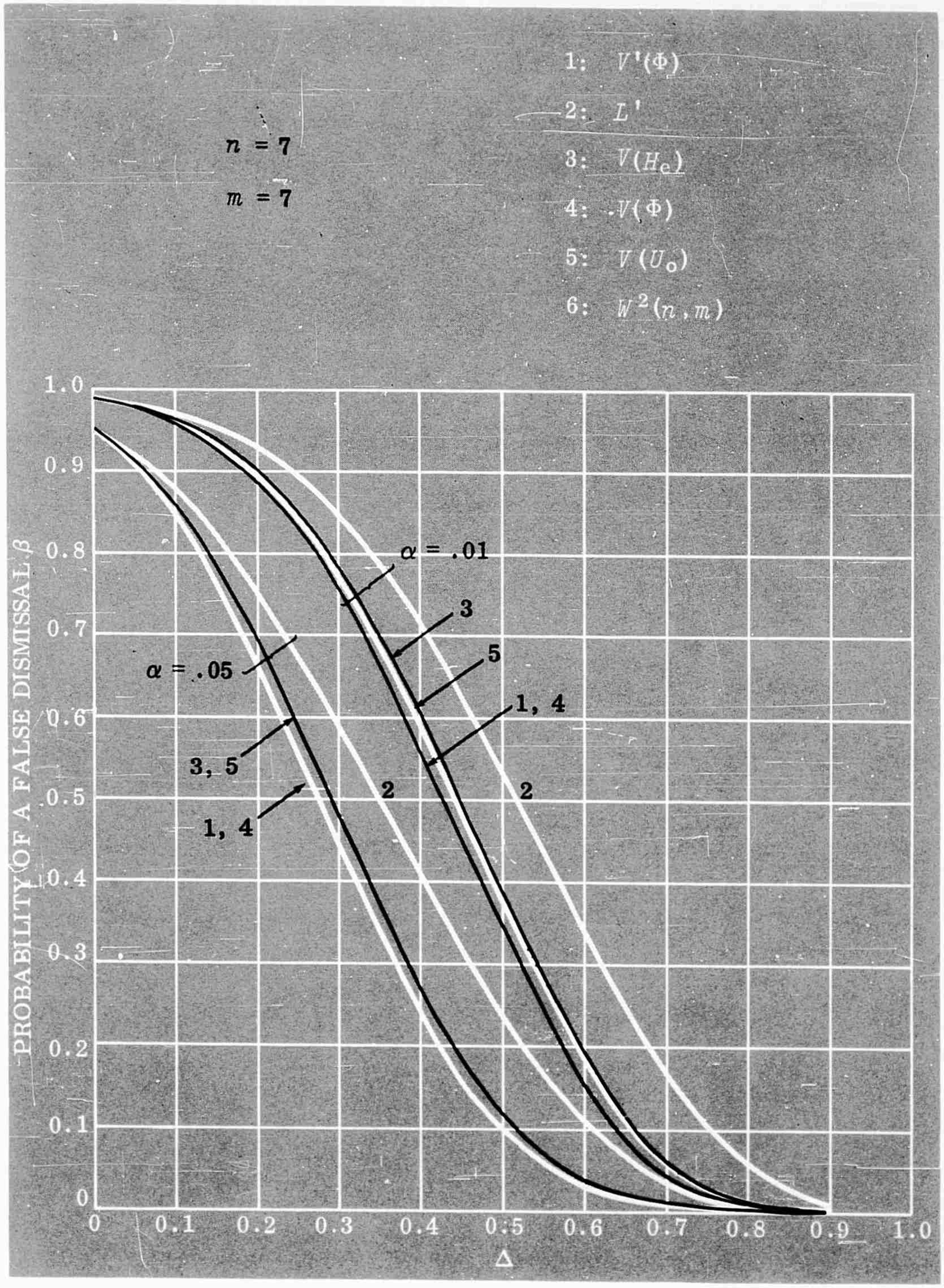
Minimum PFD for one-sided noise. (Continued)



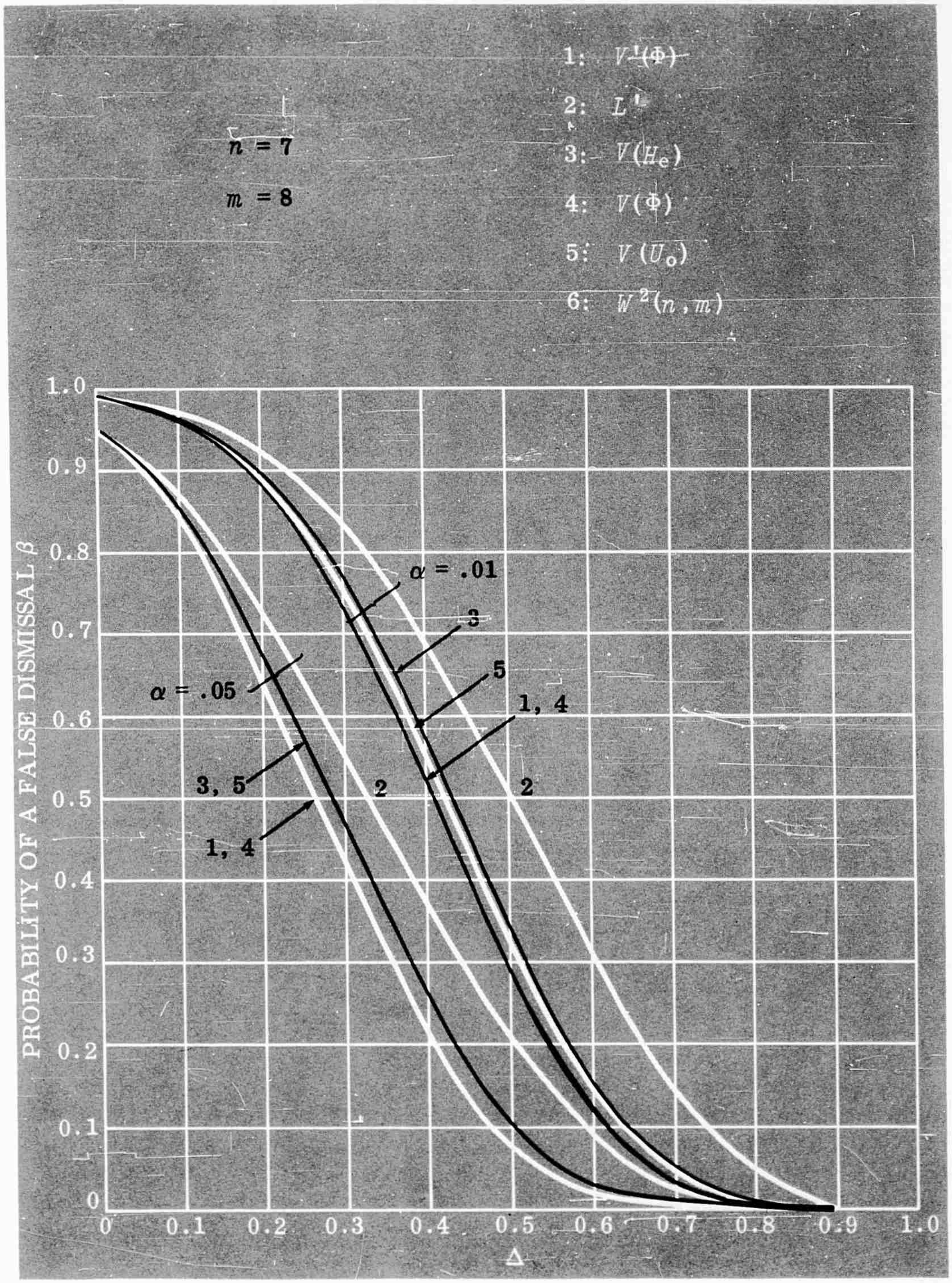
Minimum PFD for one-sided noise. (Continued)



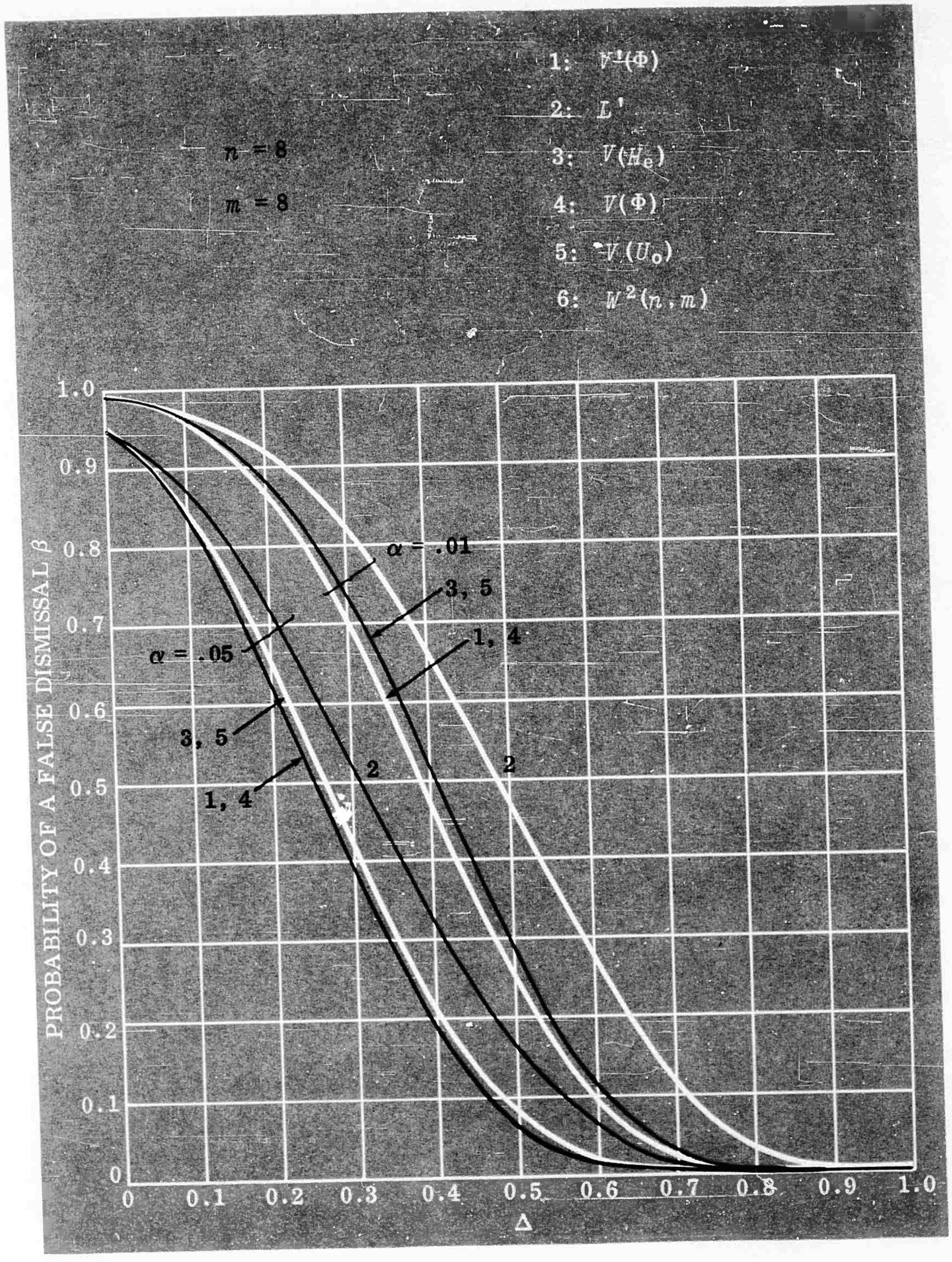
Minimum PFD for one-sided noise. (Continued)



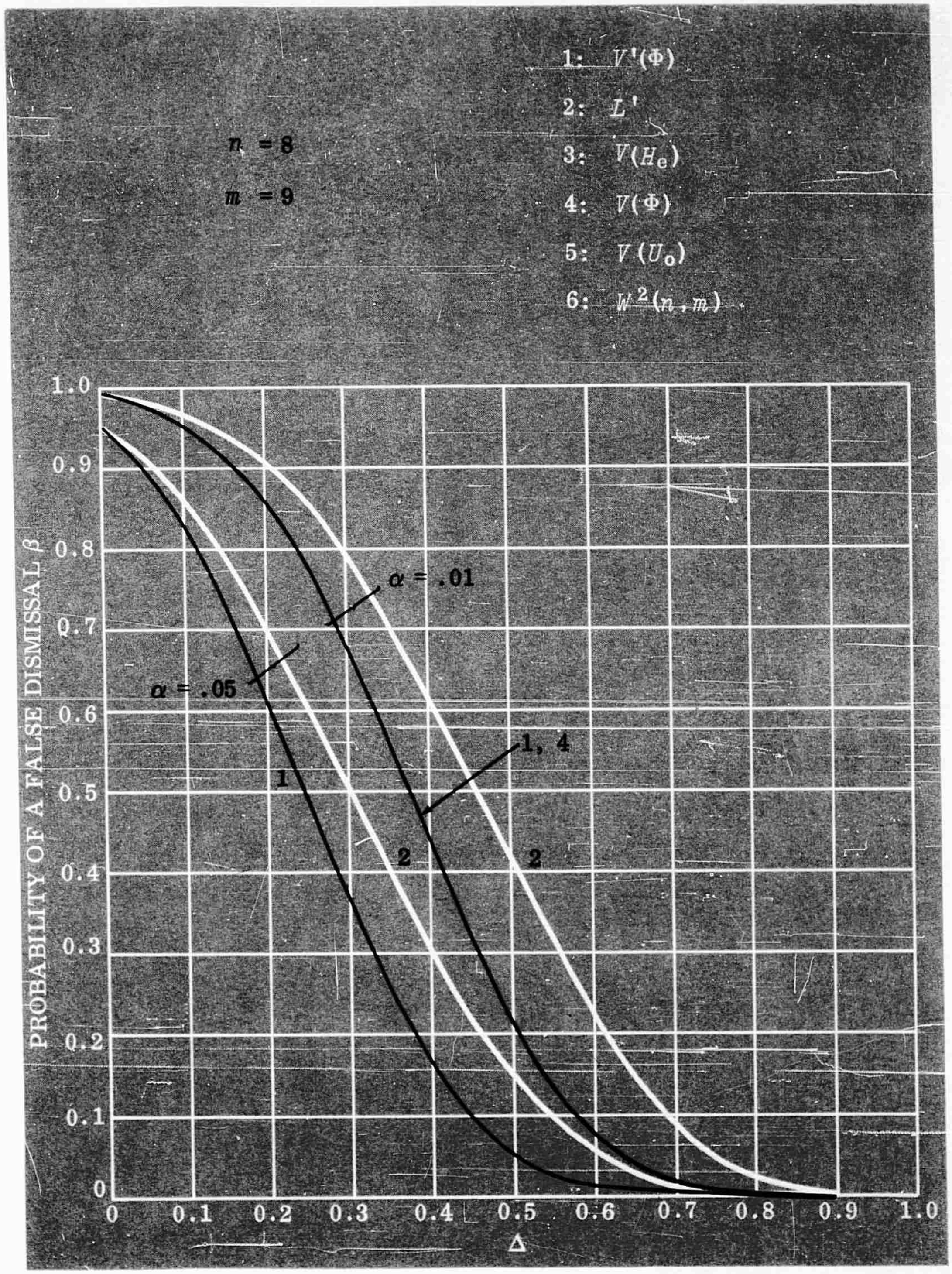
Minimum PFD for one-sided noise. (Continued)



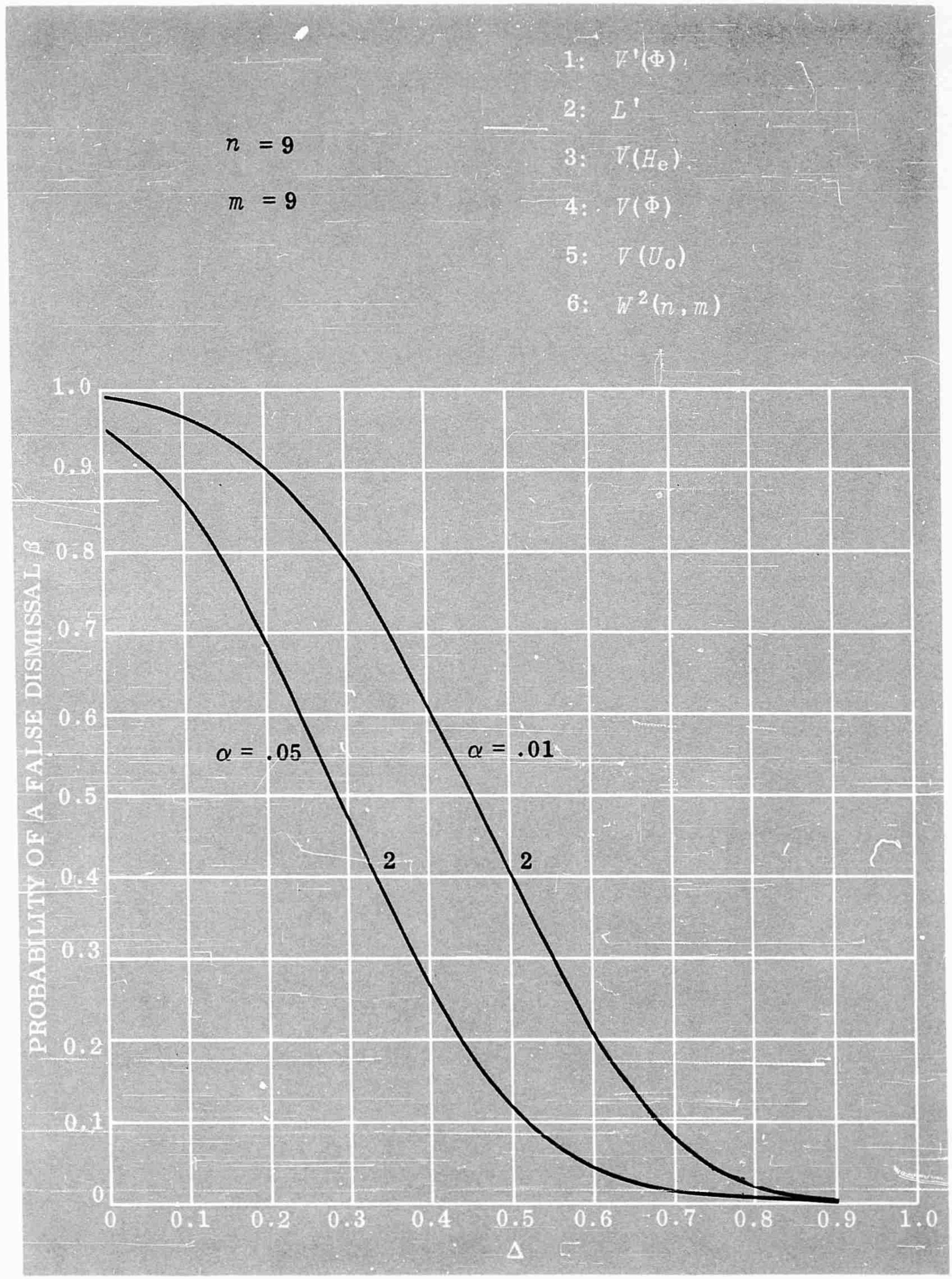
Minimum PFD for one-sided noise. (Continued)



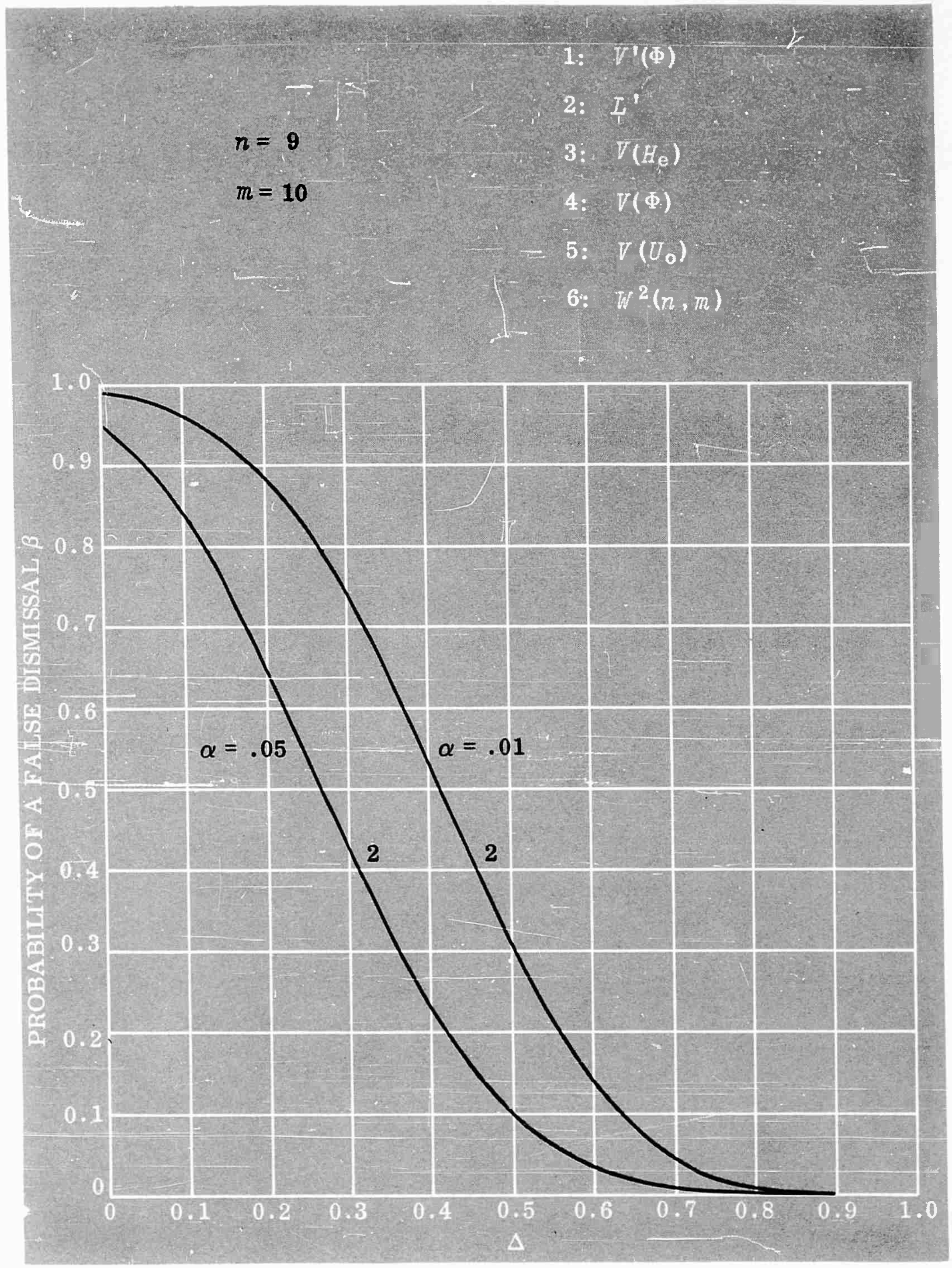
Minimum PFD for one-sided noise. (Continued)



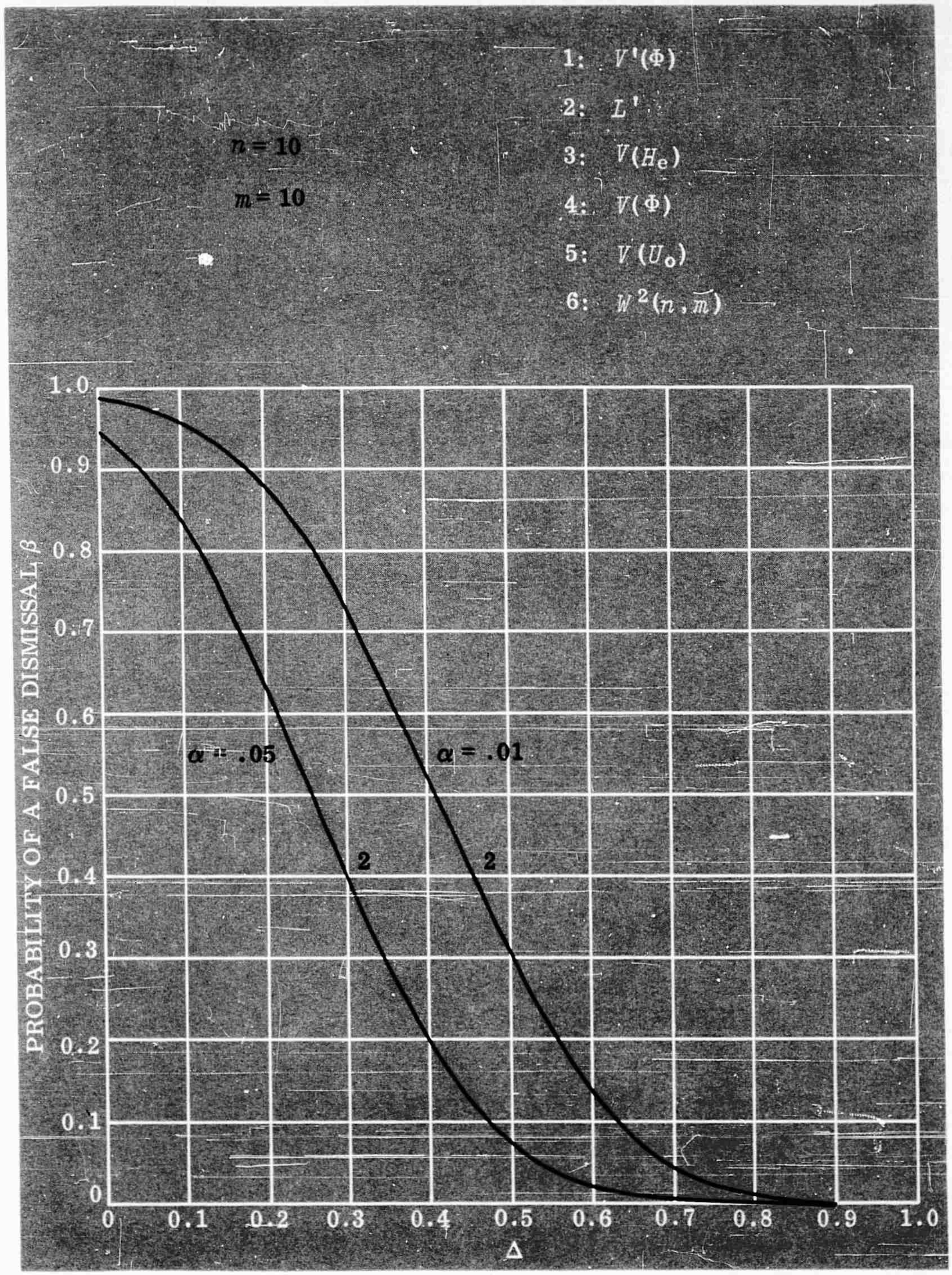
Minimum PFD for one-sided noise. (Continued)



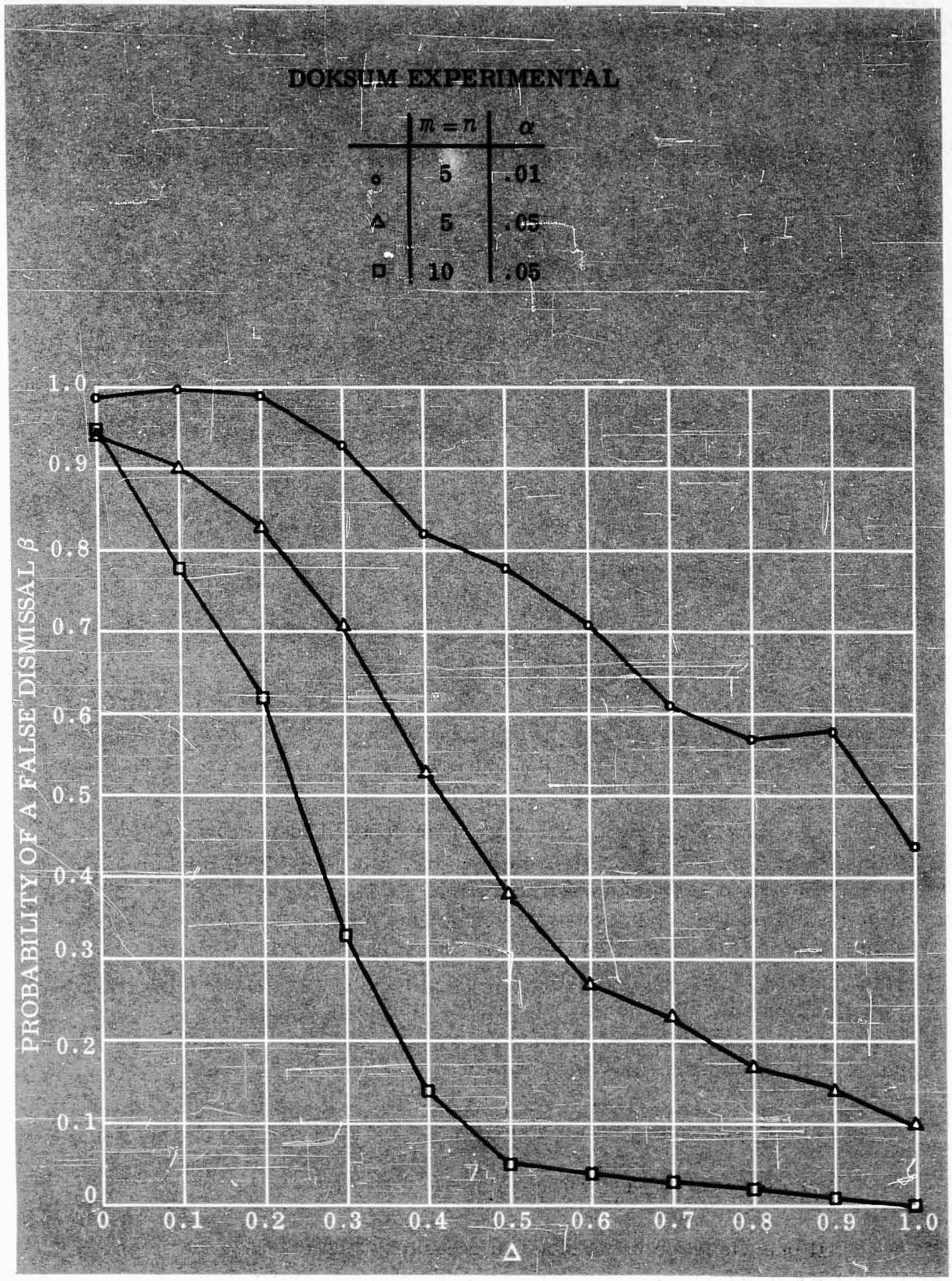
Minimum PFD for one-sided noise. (Continued)



Minimum PFD for one-sided noise. (Continued)



Minimum PFD for one-sided noise. (Continued)



Minimum PFD for one-sided noise. (Continued)

On the basis of the Bell-Moser-Thompson calculations for the Fisher-Yates, Epstein-Rosenbaum, Savage, Van de Waerden, Mann-Whitney-Wilcoxon and Cramér-von Mises Model II detectors, one concludes:

a. On the basis of minimum PFD,

$$V'(\Phi) \underset{\approx}{\}} \; V(\Phi) \underset{\sim}{\}} \; V(U_o) = V'(U_o) \} \; V'(H_e) \}\} \; L' \sim W^2(n,m)$$

where

$\underset{\approx}{\}}$ = very slightly better than

$\underset{\sim}{\}}$ = slightly better than

$\}$ = better than

$\}\}$ = much better than

$\sim$ = approximately equal

b. The max-min criterion cannot be applied here until comparable maximum PFD calculations are made.

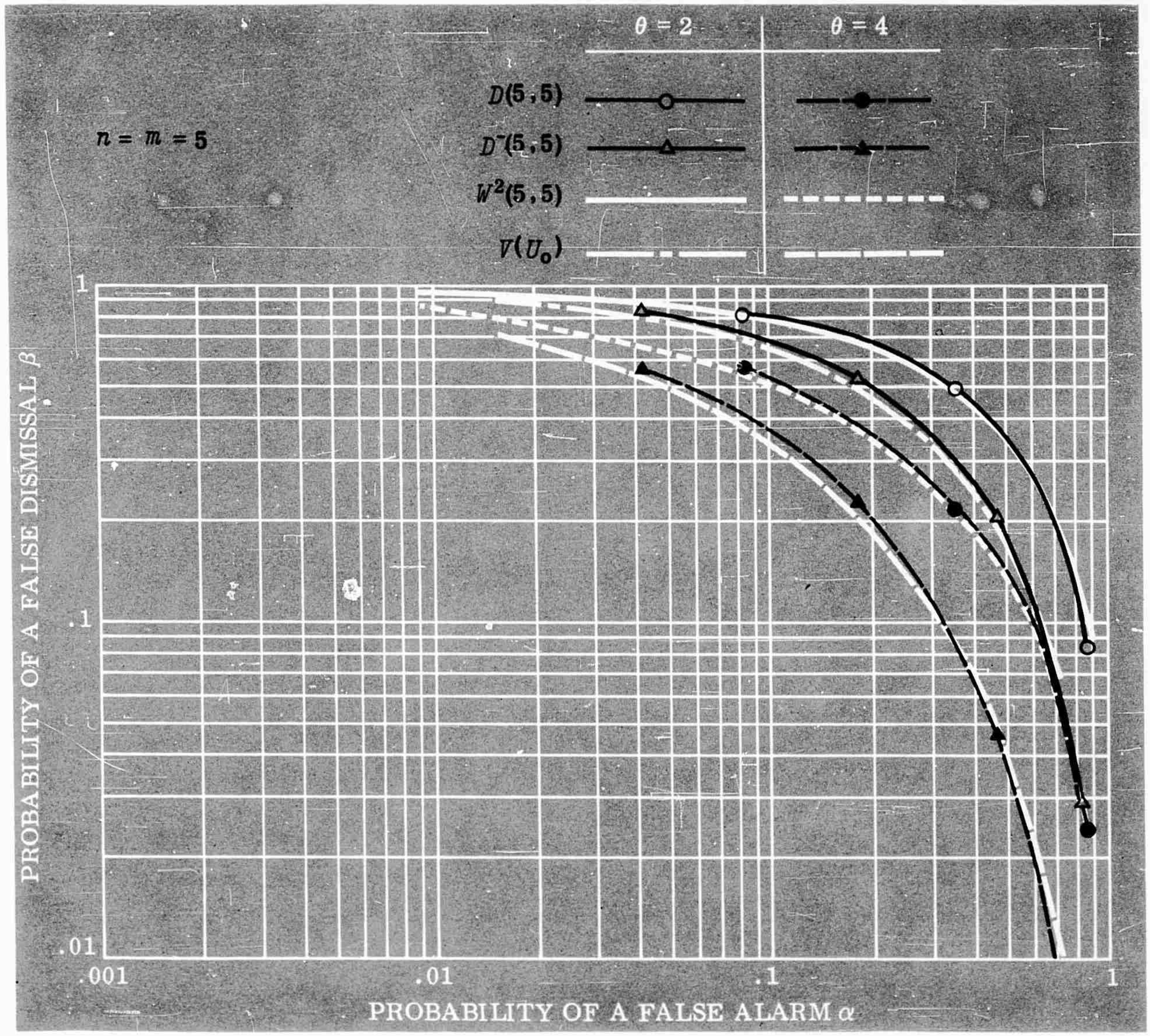
2. PFD for Rayleigh Alternatives

One next considers comparison of the Model II detectors on the basis of Rayleigh alternatives, which are considered quite important in signal detection (see illustration, p. 53). Because of time limitations, computations were made only for $D(n,m)$, $D^-(n,m)$, $W^2(n,m)$, and $V(U_o)$; for $\theta = 2$ and 4; and for sample size combinations (5, 5), (5, 10), (10, 5), and (10, 10). For computational convenience the curves are plotted in the form α vs. β for the cases mentioned in the following figures (p. 113-116). On the basis of these computations one concludes:

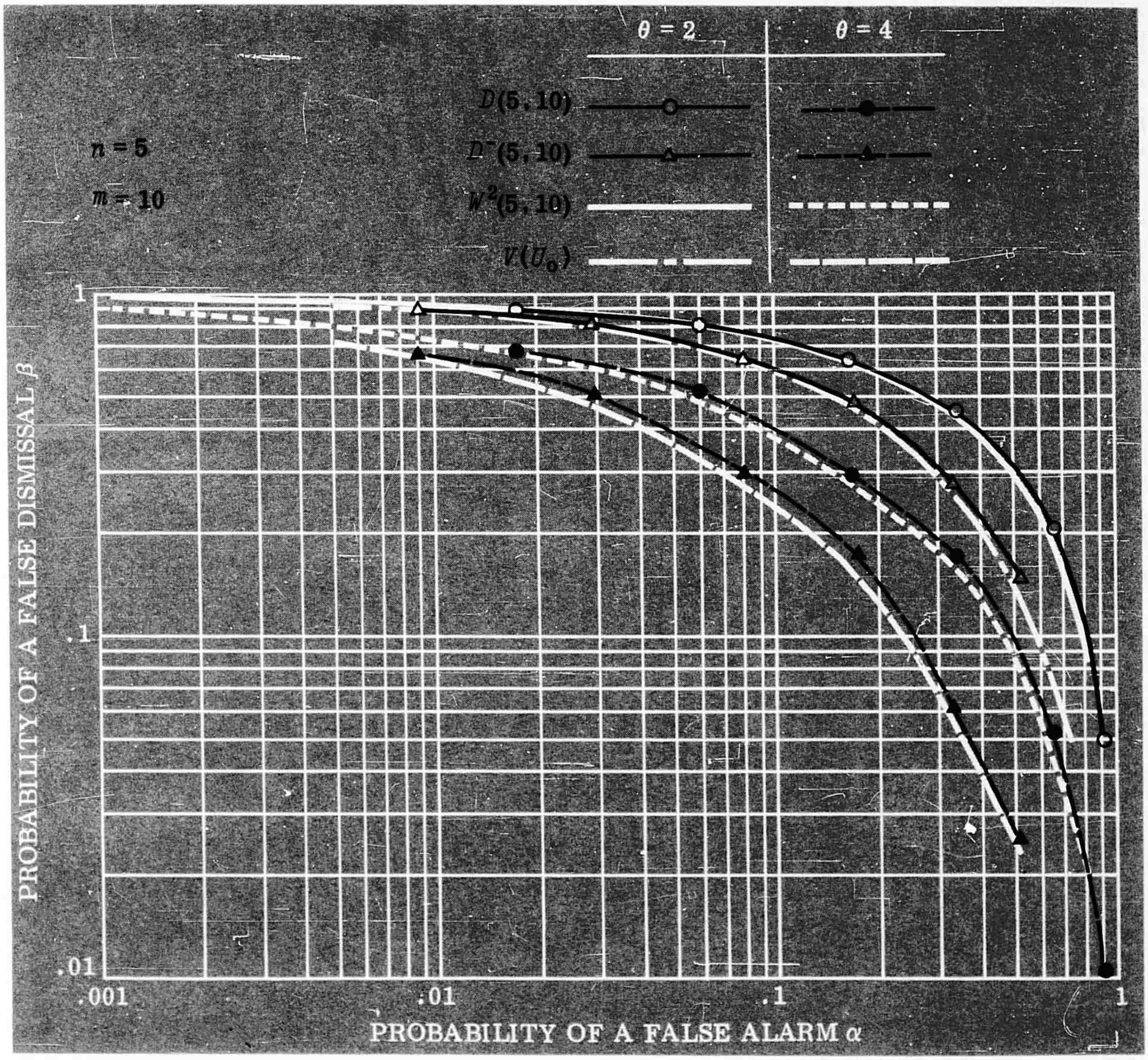
a. For Rayleigh alternatives

$$V(U_o) \underset{\sim}{\}} D^-(n,m) \}\} \; W^2(n,m) \} \; D(n,m)$$

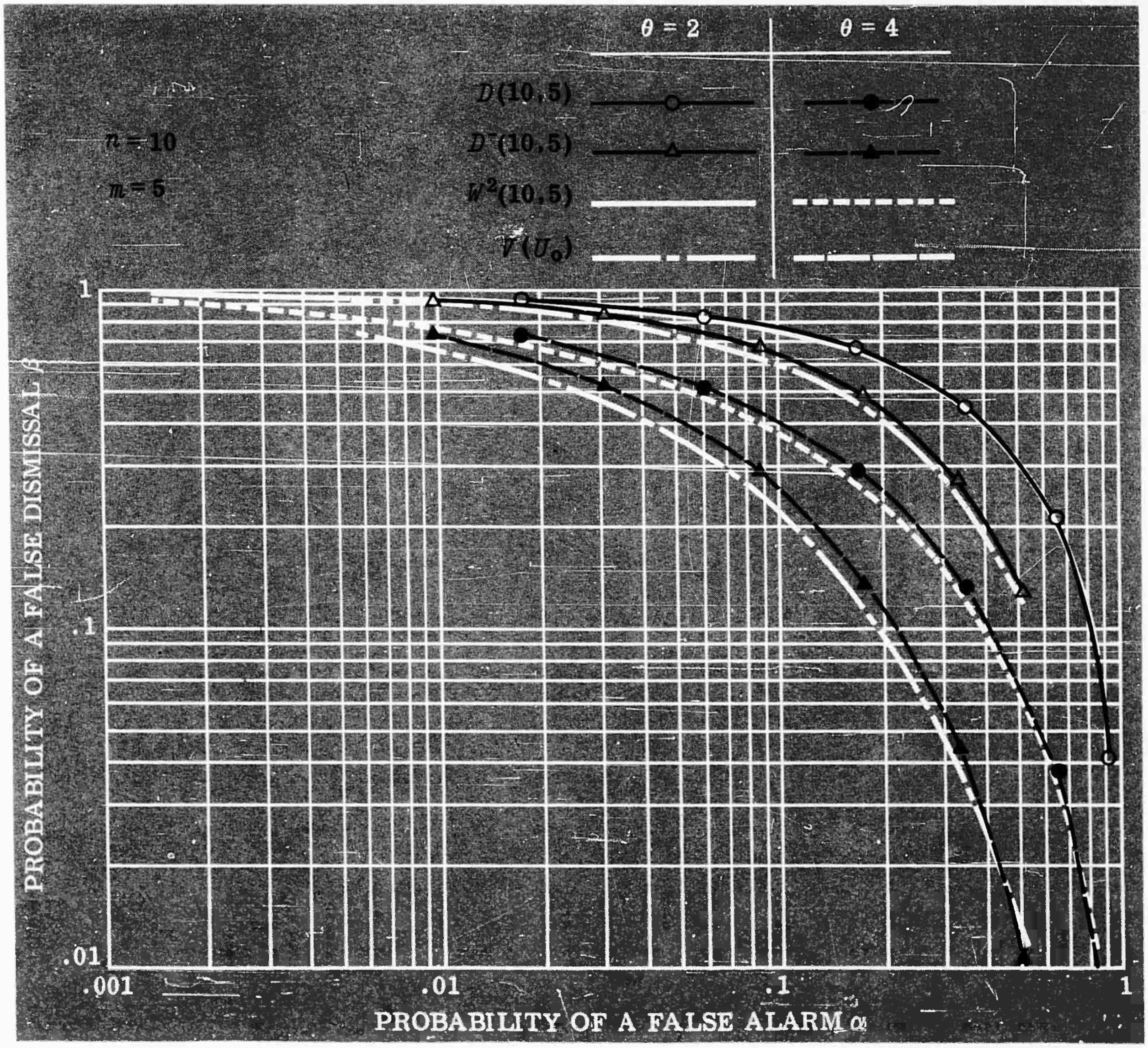
b. If one wishes to use a Model II detector which (besides its DF property) performs well against Rayleigh alternatives, one should choose a $V(U_o)$ detector, and avoid using a $D(n,m)$ detector.



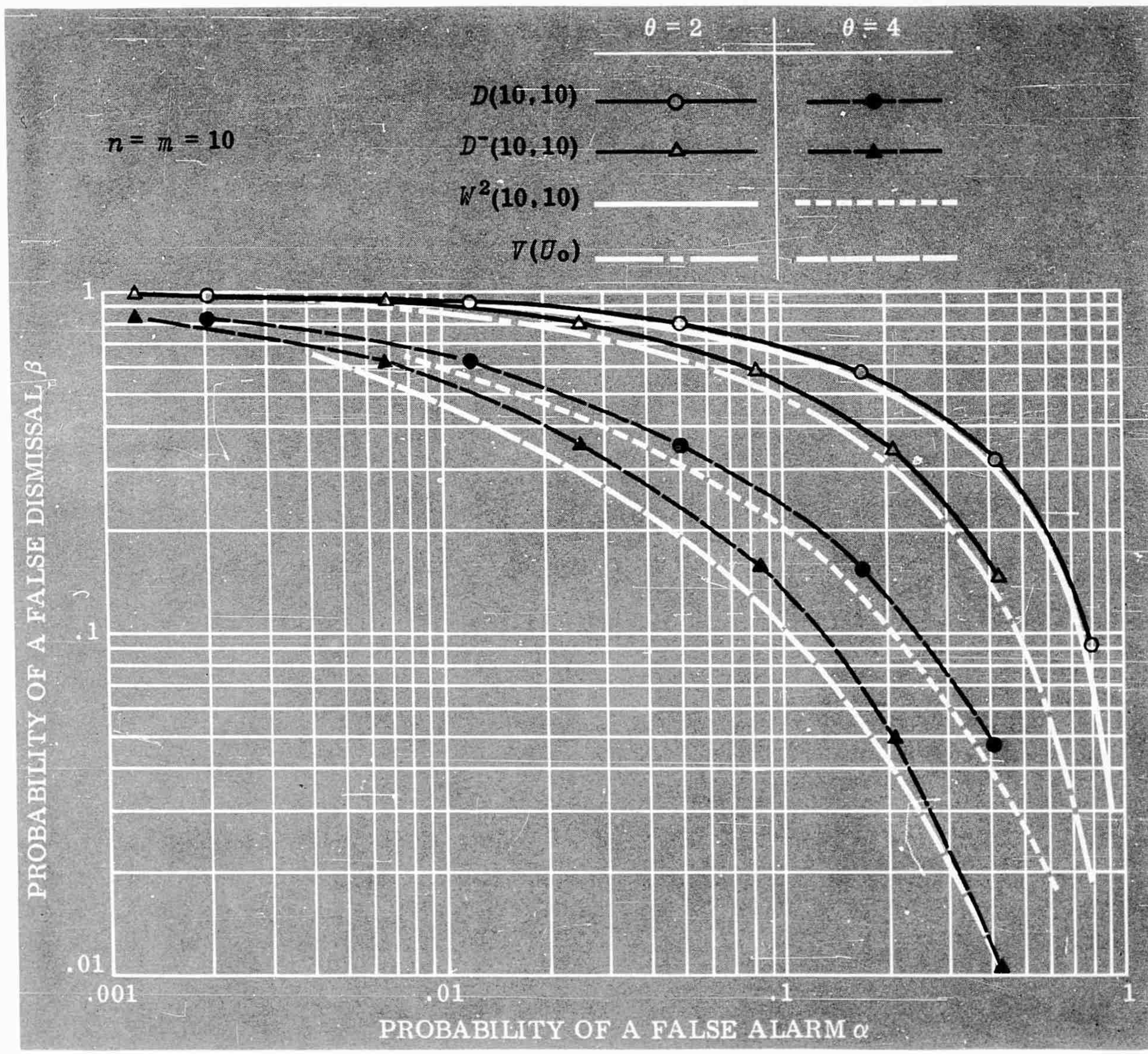
PFD vs. PFA for Rayleigh noise.



PFD vs. PFA for Rayleigh noise. (Continued)



PFD vs. PFA for Rayleigh noise. (Continued)



PFD vs. PFA for Royleigh noise. (Continued)

3. Asymptotic Relative Efficiency (ARE)

The ARE results for Model II detectors are exactly the same as those for the corresponding Model I detectors. In fact one has the following theorem from Bell and Doksum.[11]

Theorem 10. Let H be a strictly increasing continuous cpf with second moment; and let $\{G_N\}$ be a family of noise-plus-signal cpf's satisfying (a) $G_N \leq F_0$ and (b) $\lim \int G_N \, dF_0 = \frac{1}{2}$ as N tends to infinity. Then one has

$$A[V'(H), V(H)] = A[V'(H), T(H)] = A[V(H), T(H)] = 1$$

for Model II detectors; and in comparing Model I and Model II detectors,

$$A[K(H), K(H_1)] = A[V'(H), V'(H_1)] = A[V(H), V(H_1)] = A[T(H), T(H_1)] = 1$$

for all (sufficiently regular) H and H_1. Consequently, one has for translation alternatives table 14, which is an exact replica of table 6. From this table one concludes:

a. that on the basis of known ARE's there is no clear-cut ordering of the Model II detectors; but

b. that if one wishes good asymptotic performance (in the presence of increasingly weak signals) for logistic type distributions one should use a $V(U_0)$ or $T(U_0)$ detector and avoid the use of the following detectors:

$$V(H_c), \ V(\widetilde{H}_c), \ V'(H_e), \ V'(\widetilde{H}_e), \ T(H_e), \ T(\widetilde{H}_e).$$

Table 14. ARE's for Model II Detectors

TRANSLATION ALTERNATIVES \ ARE's	$A[V(\Phi), V(U_0)]$ $= A[V'(\Phi), T(U_0)]$ ...... $= A[T(\Phi), T(U_0)]$	$A[V(\Phi), V(\tilde{H}_e)]$ $= A[V'(\Phi), T(\tilde{H}_e)]$ ...... $= A[T(\Phi), T(\tilde{H}_e)]$	$A[V(\Phi), V(H_e)]$ $= A[V'(\Phi), T(H_e)]$ ...... $= A[T(\Phi), T(H_e)]$	$A[V(\tilde{H}_e), V(U_0)]$ $= A[V'(\tilde{H}_e), T(U_0)]$ ...... $= A[T(\tilde{H}_e), T(U_0)]$	$A[V(H_e), V(U_0)]$ $= A[V'(H_e), T(U_0)]$ ...... $= A[T(H_e), T(U_0)]$	$A[V(H_e), V(\tilde{H}_e)]$ $= A[V'(H_e), T(\tilde{H}_e)]$ ...... $= A[T(H_e), T(\tilde{H}_e)]$
Normal $F_o(x) = \Phi(x)$	1.05					
Uniform $F_o(x) = U_o(x)$	∞			∞	∞	1.00
Exponential $F_o(x) = H_e(x)$	∞	∞		.333		
Negative exponential $F_o(x) = \tilde{H}_e(x)$ $= e^x,\ x \leq 0$	∞		∞		.333	
Logistic $F_o(x) = \dfrac{1}{1 + e^{-x}}$	.955	1.27	1.27	.75	.75	1.00
Double exponential $F_o(x) = \begin{cases} e^x/2, & x \leq 0 \\ 1 - e^x/2, & x > 0 \end{cases}$	.85					
Cauchy $F_o(x) = \dfrac{1}{2} + \dfrac{1}{\pi}\tan^{-1} x$	.71					

4. Almost Locally Minimum (ALM) PFD Model II Detectors

Using the randomized statistics of Section III D and the following theorem of Bell and Doksum[12], one can construct ALM PFD detectors for certain specified noise classes.

Theorem 11. If conditions (i), (ii), (iii) and one of (iva), (ivb) and (ivc) below hold for strictly increasing continuous pure-noise cpf F_o and the noise-plus-signal class $\{G_\theta = h_\theta{}' (F_o)\}$, then a Model II detector which bases its YES-NO decisions on the statistic

$$T(h, \theta) = \frac{1}{n} \sum_{i=1}^{n} J_\theta \{U[R(X_i)]\} - \frac{1}{m} \sum_{j=1}^{m} J_\theta \{U[R(Y_j)]\}$$

or on the statistic

$$T'(h, \theta_o) = \frac{1}{n} \sum_{i=1}^{n} J_o{}' \{U[R(X_i)]\} - \frac{1}{m} \sum_{j=1}^{m} J_o{}' \{U[R(Y_j)]\}$$

(where $J_\theta(u) = \ln h_o{}'(u)$; $J_o{}'(u) = \frac{\partial}{\partial\theta} \ln h_\theta{}'(u) \big|_{\theta=\theta_o}$ and $U(1), \cdots, U(N)$ are the order statistics of a U_o-random noise generator) is ALM PFD.

Regularity Conditions:

(i) h_θ has a derivative $h_\theta{}'$, which, for almost all u, is continuous in some nondegenerate closed interval $I(\theta_o)$ containing θ_o, and satisfies $h_{\theta_o}{}'(u) \equiv 1$;

(ii) there exist functions $M_o(u)$ and $M_1(u)$ integrable over (0, 1) and such that for all u and all θ in $I(\theta_o)$,

$$0 < h_\theta{}'(u) \leq M_o(u) \text{ and } |h_\theta(u)| \leq M_1(u);$$

(iii) $0 < \lim \frac{m}{n} = r < \infty$ as N tends to infinity;

(iva) J_0' is the inverse of a cpf H with finite second moment;

(ivb) J_0' is convex, $\int_0^1 [J_0'(u)]^2 \, du < \infty$ and J'_0 is bounded below; and

(ivc) there exists a continuous cpf H such that $J_0'[H(x)]$ is convex, bounded from below, $\left| \int x dH(x) \right| < \infty$, and $\int J_0'[H(x)] \, dH(x) < \infty$.

Table 15 gives the J_θ and J_0' for detectors which are ALM PFD with respect to some special classes of noise-plus-signal cpf's. Since the detectors in question are SDF, the functions J_0 and J_0' depend only on $h_\theta = F_\theta F_0^{-1}$.

Table 15. ALM PFD Detectors for Certain Noise-plus-signal Classes $\{F_o(x)\}$ of epf's

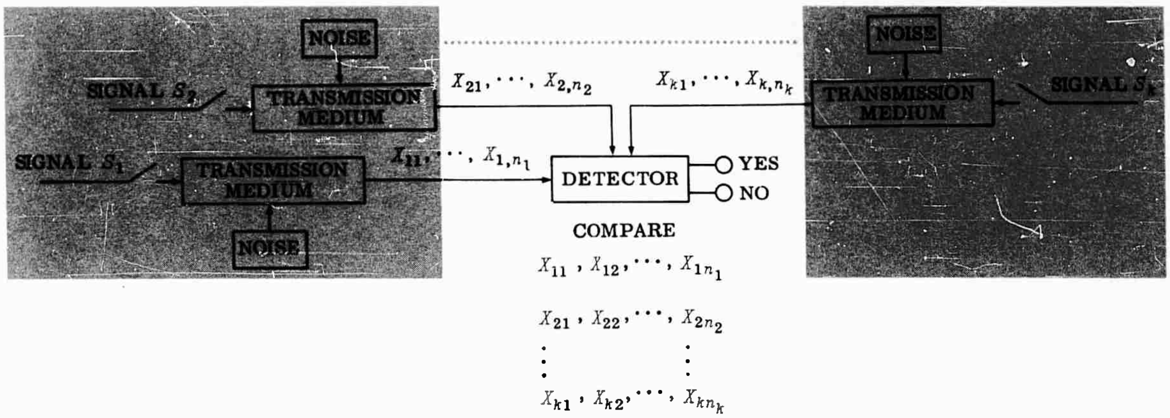
[The statistics of the ALM PFD detectors are of the form $S = \frac{1}{n} \sum_{i=1}^{n} J\{U[R(X_i)]\} - \frac{1}{m} \sum_{j=1}^{m} J\{U[R(Y_j)]\}$

where $J = J_\theta$ or J_o' as given below]

$h_\theta(u)$	θ_o	$a\ J_\theta(u) + b$ (where a, b, c and d are used for convenience)	$c\ J_o'(u) + d$	General Example	Specific Example
u^θ	1	$\ln u$	$\ln u$	$F_o(x)$ vs. $F_o^{\theta}(x)$	
$1-(1-u)^\theta$	1	$\ln(1-u)$	$\ln(1-u)$	$F_o(x)$ vs. $\{1-[1-F_o(x)]^\theta\}$	$1-e^{-x}$ vs. $1-e^{-x\theta}$
$(1-\theta)u+\theta u^2$	0	$\ln(1-\theta+2\theta u)$	u	$F_o(x)$ vs. $[(1-\theta)F_o(x)+\theta F_o^2(x)]$	
$(1-\theta)u+\theta u^{K+1}$	0	$\ln[1-\theta+(K+1)\theta u^{K+1}]$	u^{K+1}	$F_o(x)$ vs. $[(1-\theta)F_o(x)+\theta F_o^{K+1}(x)]$	
$\frac{e^{\theta u}-1}{e^\theta-1}$	0	u	u	$F_o(x)$ vs. $\frac{e^{\theta F_o(x)}-1}{e^\theta-1}$	
$\frac{u}{u+e^\theta(1-u)}$	0	$\ln[u+e^\theta(1-u)]$	u	$F_o(x)$ vs. $\frac{F_o(x)}{F_o(x)+e^\theta[1-F_o(x)]}$	$\frac{1}{1-e^{-x}}$ vs. $\frac{1}{1+e^{-x+\theta}}$
$G[G^{-1}(u)-\theta]$	0	$\ln\{G'[G^{-1}(u)-\theta]\}$ $-\ln\{G'[G^{-1}(u)]\}$	$\frac{G''[G^{-1}(u)]}{G'[G^{-1}(u)]}$	$F_o(x)$ vs. $F_o(x-\theta)$	above; and $\Phi(x)$ vs. $\Phi(x-\theta)$
$G\left[\frac{G^{-1}(u)}{\theta}\right]$	1	$\ln\left\{G'\left[\frac{G^{-1}(u)}{\theta}\right]\right\}$ $-\ln\{G'[G^{-1}(u)]\}$	$\frac{G''[G^{-1}(u)]\ G^{-1}(u)}{G'[G^{-1}(u)]}$	$F_o(x)$ vs. $F_o(x/\theta)$	$\Phi(x)$ vs. $\Phi(x/\theta)$
$\exp\{a(\theta)J(\theta)+b(\theta)+Q(u,\theta)\}$ $= h_\theta'(u)$ where $Q(u,\theta)=0(\theta-\theta_o)$	θ_o	$J_\theta = J_o'$	$J_\theta = J_o'$	$F_o(x)$ vs. $\int h_\theta[F_o(x)]\ F_o'(x)\,dx$	

Table 16. Statistical Distributions for Model II Detectors

Statistic	Distribution	Source of Tables, Formulas, etc. (numbers refer to Bibliography)
$D(n,m)$	(2-sample Kolmogorov-Smirnov)	21, p, 434 ff
$D^-(n,m)$, $D^+(n,m)$	(2-sample, 1-sided Kolmogorov-Smirnov)	21, p. 431 ff
$W^2(n, m)$	(2-sample Cramér-von Mises)	21, p. 443 ff; 23
$W^+(n,m)$, $W^-(n,m)$	[Equivalent to $V(U_o)$]	
Q_1	Asymptotically chi-square (median)	21, p. 49 ff; 28, p. 226 ff
Q'_{s+1}	Asymptotically chi-square (extended median)	21, p. 49 ff; 28, p. 296 ff
$\hat{Q}_{s+1}$	Asymptotically chi-square (Matusita)	21, p. 49 ff; 25
$\mathcal{E}$	(Empty block)	29, p. 446
$S(n, m)$	(2-sampie Sherman)	
$S_2(n, m)$	(2-sample Kimball-Moran)	
L'	(Epstein-Rosenbaum)	21, 499 ff
L	(Moses "extreme reaction")	30
$V(U_o)$, $V''(U_o)$	(Mann-Whitney-Wilcoxon)	22, p. 325 ff
$V''(\Phi)$	Asymptotically normal (Fisher-Yates)	8, p. 498; 31, p. 66
$V(\Phi)$	Asymptotically normal (Van de Waerden)	32
$T(\Phi)$	Normal (Doksum)	11, 12



IV. MODEL III DETECTORS

As mentioned in Section I, the Model III detectors are extensions of the Model II detectors. The region being searched or examined is divided into k subsectors and the PS data from the k subsectors are compared. The detector, then, on the basis of the appropriate k-sample statistic, decides NO iff there is no significant difference between the k PS data samples:

Sector 1: $X_{11}, X_{12}, \cdots, X_{1,n_1}$

Sector 2: $X_{21}, X_{22}, \cdots, X_{2,n_2}$

$\cdot \quad \cdot \quad \cdot$

$\cdot \quad \cdot \quad \cdot$

$\cdot \quad \cdot \quad \cdot$

Sector k: $X_{k1}, X_{k2}, \cdots, X_{k,n_k}$

The three "natural" subclasses of Model III detectors are:

(1) Sample cpf (SDF Scheffé Model III) detectors;

(2) Run-block (SDF Scheffé Model III) detectors; and

(3) Rank-sum (SDF Scheffé Model III) detectors.

These three subclasses are introduced and illustrated in the next section, where the following notation will be employed

n_i = the number of data points in the i^{th} PS data sample;

k = number of sectors; and $N = \sum_{i=1}^{k} n_i$.

PS sample from sector i: $X_{i1}, X_{i2}, \cdots, X_{i,n_i}$

Order statistics: $X_i(1) < X_i(2) < \cdots < X_i(n_i)$

Combined sample: $\{Z_1, \cdots, Z_N\} = \{X_{11}, \cdots, X_{1,n_1}; \cdots; X_{k1}, \cdots, X_{k,n_k}\}$

Order statistics: $Z(1) < \cdots < Z(N)$

Degenerate cpf ϵ, where $\epsilon(x) = 0$ or 1 according as $x < 0$ or $x \geq 0$.

Sample cpf of i^{th} sample: $\bar{F}_i$, where

$$\bar{F}_i(x) = \frac{1}{n_i} \sum_{j=1}^{n_i} \epsilon(x - X_{ij})$$

Combined sample cpf: $\bar{F}$, where

$$\bar{F}(Z) = \frac{1}{N} \sum_{r=1}^{N} \epsilon(z - Z_r)$$

Rank of X_{ij} in combined sample: $R(X_{ij})$

$$R_{i+} = \sum_{j=1}^{n_i} R(X_{ij}),$$ the sum of the ranks of the members of the i^{th} sample

Artificial noise: $W_1, \cdots, W_N$

Order statistics: $W(1) < \cdots < W(N)$

$$W_{i+} = \sum_{j=1}^{n_i} W[R(X_{ij})],$$ the sum of the artificial noise values having the same ranks as the members of the i^{th} sample

$$W_{++} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} W[R(X_{ij})],$$ the sum of all of the artificial noise values

H: the cpf of the artificial noise sample

$$E_{i+}(H) = \sum_{j=1}^{n_i} E\left\{W[R(X_{ij})] \mid H\right\}$$

$$E_{++}(H) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} E\left\{W[R(X_{ij})] \mid H\right\}$$

$$H_{i+}^{-1} = \sum_{j=1}^{n_i} H^{-1}\left(\frac{R(X_{ij})}{N+1}\right)$$

$$H_{++}^{-1} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} H^{-1}\left(\frac{R(X_{ij})}{N+1}\right)$$

For some of the detectors to be described it will be necessary to consider blocks of the order statistics $Z(1) < \cdots < Z(N)$. The blocks are chosen to be of equal size (if possible), and, consequently (wherever possible),

b, the number of blocks, is a positive integer which divides N and

The b blocks of order statistics are, then,

$$\left[Z(1), \cdots, Z\left(\frac{N}{b}\right)\right], \left[Z\left(\frac{N}{b}+1\right), \cdots, Z\left(\frac{2N}{b}\right)\right], \cdots,$$

$$\left[Z\left(N-\frac{N}{b}+1\right), \cdots, Z(N)\right].$$

n_{ij} = number of members of the i^{th} sample in the j^{th} block

$$\left[Z\left(\frac{jN}{b}-\frac{N}{b}+1\right), \cdots, Z\left(\frac{jN}{b}\right)\right]$$

Further, it will be helpful to recall some of the more useful relations between these quantities

$$R(X_{ij}) = \sum_{r=1}^{N} \epsilon(X_{ij} - Z_r) = N\,\bar{F}(X_{ij})$$

$$\bar{F}(z) = \frac{1}{N}\sum_{i=1}^{k} n_i\,\bar{F}_i(z), \quad \bar{F}(Z(r)) = \frac{r}{N}$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i} R(X_{ij}) = \frac{1}{2}N(N+1) = \sum_{i=1}^{k} R_{i+}$$

$$\sum_{i=1}^{k} W_{i+} = W_{++}; \quad \sum_{i=1}^{k} E_{i+} = E_{++}, \quad \sum_{i=1}^{k} H_{i+}^{-1} = H_{++}^{-1}$$

$$\sum_{i=1}^{k} n_i\left(\frac{W_{i+}}{n_i} - \frac{W_{++}}{N}\right)^2 = \sum_{i=1}^{k}\frac{W_{i+}^{\,2}}{n_i} - \frac{W_{++}^{\,2}}{N}$$

$$\sum_{j=1}^{b} n_{ij} = n_i, \quad \sum_{i=1}^{k} n_{ij} = N/b$$

One can now introduce and illustrate the model III detectors.

A. MODEL III SAMPLE Cpf DETECTORS

The sample cpf detectors are those based on the given statistics.

1. Kolmogorov-Smirnov Detector

$$D(n_1, \cdots, n_k ; s ; \Psi)$$

$$= \max_r \sum_{i=1}^{k} \left\{ \left| \bar{F}_i (Z(r)) - \bar{F}(Z(r)) \right| \sqrt{n_i}\, \Psi[\bar{F}(Z(r))] \right\}^s , \tag{68}$$

for which the most usual special case is

$$D(n_1, \cdots, n_k ; 2; 1) = \max_r \left\{ \sum_{i=1}^{k} n_i \; \bar{F}_i^{\,2}(Z(r)) - \frac{r^2}{N} \right\} \tag{69}$$

2. Cramér-von Mises Detector

$$W(n_1, \cdots, n_k ; s ; \Psi)$$

$$= \frac{1}{N} \sum_{r=1}^{N} \sum_{i=1}^{k} \left\{ \left| \bar{F}_i (Z(r)) - \bar{F}(Z(r)) \right| \sqrt{n_i}\, \Psi[\bar{F}(Z(r))] \right\}^s \tag{70}$$

for which the usual special case is

$$W^2(n_1, \cdots, n_k ; 2; 1) = \frac{1}{N} \sum_{r=1}^{N} \sum_{i=1}^{k} n_i \; \bar{F}_i^{\,2}(Z(r)) - \frac{(N+1)(2N+1)}{6N} \tag{71}$$

3. Sign-Quantile Detector

$$Q(n_1, \cdots, n_k ; q_1, \cdots, q_m ; s, \Psi)$$

$$= \sum_{t=1}^{m+1} \sum_{i=1}^{k} \left[\left| \bar{\bar{F}}_i \, (Z(Nq_t)) - \bar{\bar{F}}_i \, (Z(Nq_{t-1})) - (q_t - q_{t-1}) \right| \Psi_{i,t} \right]^s \qquad (72)$$

for which the most usual version is

$$Q(n_1, \cdots, n_k ; b) = b \sum_{i=1}^{k} \sum_{j=1}^{b} \frac{n_{ij}^2}{n_i} - N \qquad (73)$$

B. MODEL III RUN-BLOCK DETECTORS

Model III run-block detectors are used very sparingly, but the most usual one is based on the Mosteller-Tukey statistic

$$\hat{M}, \text{ length of last run, } = \max_{i,t} \; n_i \left[1 - \bar{F}_i \left(X_t (n_t) \right) \right] \qquad (74)$$

C. MODEL III RANK-SUM DETECTORS

Model III rank-sum detectors are based on statistics corresponding to the numerator of the "F-statistic" in the ordinary analysis of variance

$$S(H) = \sum_{i=1}^{k} n_i \left(\frac{H_{i+}^{-1}}{n_i} - \frac{H_{++}^{-1}}{N} \right)^2 = \sum_{i=1}^{k} \frac{(H_{i+}^{-1})^2}{n_i} - \frac{H_{++}^2}{N} \qquad (75)$$

Usually one chooses $H = U_o$ to obtain the statistic of the Kruskal-Wallis detector

$$S(U_o) = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_{i+}^2}{n_i} - 3(N+1) \tag{76}$$

or $H = \Phi$ to obtain the Model III Van de Waerder detector

$$S(\Phi) = \sum_{i=1}^{k} \frac{(H_{i+}^{-1})^2}{n_i} \tag{77}$$

The detectors based on the expected values $E[W(r)|H]$ instead of the percentiles $H^{-1}\left(\frac{r}{N+1}\right)$ are those whose statistics are of the form

$$S'(H) = \sum_{i=1}^{k} n_i \left(\frac{E_{i+}(H)}{n_i} - \frac{E_{++}(H)}{N}\right)^2 = \sum_{i=1}^{k} \frac{E_{i+}^2(H)}{n_i} - \frac{E_{++}^2(H)}{N} \tag{78}$$

$$S'(U_o) = S(U_o), \text{ and} \tag{79}$$

$$S'(\Phi) = \sum_{i=1}^{k} \frac{E_{i+}^2(H)}{n_i} \tag{80}$$

are the usual special cases of this statistic.

D. MODEL III ARTIFICIAL NOISE DETECTORS

The final Model III rank-sum detectors are those which make use of artificial noise in making decisions. (See illustrations, pp. 90, 96). The statistics employed here are based on an idea of Doksum[11] and are of the form

$$\hat{S}(H) = \sum_{i=1}^{k} n_i \left(\frac{W_{i+}}{n_i} - \frac{W_{++}}{N} \right)^2 = \sum_{i=1}^{k} \frac{W_{i+}^2}{n_i} - \frac{W_{++}^2}{N} \qquad (81)$$

where H is the cpf of the artificial noise random sample $W_1, \cdots, W_N$.

$\hat{S}(\Phi)$ is the most common version in use. (82)

To illustrate the computations involved, the following example is presented.

Example 13. The region being searched is divided into four subregions, and the PFA is set at $\alpha = .05$. The data from the sectors turn out to be

Sector 1: 4.7, 4.1, 4.5, 7.6, 6.9, 5.8

Sector 2: 4.8, 4.4, 4.9, 5.7, 7.0

Sector 3: 4.3, 5.4, 3.4, 4.6, 5.0

Sector 4: 5.6, 5.3, 6.4, 6.0, 5.2, 6.5, 5.1, 4.2

Consequently, one has

k=4, n_1=6, n_2=5, n_3=5, n_4=8, and N=24.

The decision rules for the various detectors are as given.

$\hat{M}$ detector: Decide YES iff $\hat{M} \geq 4$, the 95th percentile of the Mosteller-Tukey table.[33]

$W^2(6, 5, 5, 8)$ detector: Decide YES iff $W^2 > 1.00$, the 95th percentile of the Kiefer table.[34]

$D(6, 5, 5, 8)$ detector: Decide YES iff $D > 3.06$, the 95th percentile of the Kiefer table.[33]

$S(U_0) = S'(U_0)$ detector: Decide YES iff $S(U_0) > 7.81$, the 95th percentile of chi-square distributions with 3 degrees of freedom.

$S(\Phi)$ detector: Decide YES iff $S(\Phi) > 7.81$.

$Q(6, 5, 5, 8; 3)$ detector: Decide YES iff $Q > 12.59$, the 95th percentile of the chi-square distribution with 6 degrees of freedom.

Since all of the detectors being considered base their decisions solely on functions of the ranks $R(X_{ij})$, the data may be ranked as shown in table 17.

The last column of the table contains the ordered values of the data produced by an artificial normal noise generator: 0.344, -0.664, 1.351, -0.429, -2.510, -0.148, -0.132, -0.605, 0.379, 0.394, 0.526, -1.354, 1.119, 0.705, -1.167, 0.256, -0.517, -0.084, 1.553, 0.588, 0.336, 0.220, -0.177

r	$Z(r)$	Sample	$\bar{F}_1(Z(r))$	$\bar{F}_2(Z(r))$	$\bar{F}_3(Z(r))$	$\bar{F}_4(Z(r))$	$\sum n_i \bar{F}_i^2(Z(r)) = S_+(r)$	$S_+(r) - \frac{r^2}{N}$	Artificial Normal Noise $W(r)$
1	3.4	3	0	0	.200	0	.200	.158	-2.510
2	4.1	1	.167	0	.200	0	.367	.200	-1.354
3	4.2	4	.167	0	.200	.125	.492	.117	-1.167
4	4.3	3	.167	0	.400	.125	1.092	.425	-0.664
5	4.4	2	.167	.200	.400	.125	1.292	.250	-0.605
6	4.5	1	.333	.200	.400	.125	1.791	.291	-0.517
7	4.6	3	.333	.200	.600	.125	2.791	.749	-0.429
8	4.7	1	.500	.200	.600	.125	3.625	.958	-0.177
9	4.8	2	.500	.400	.600	.125	4.225	.850	-0.148
10	4.9	2	.500	.600	.600	.125	5.225	1.058	-0.132
11	5.0	3	.500	.600	.800	.125	6.625	1.583	-0.084
12	5.1	4	.500	.600	.800	.250	7.000	1.000	0.014
13	5.2	4	.500	.600	.800	.375	7.625	.583	0.220
14	5.3	4	.500	.600	.800	.500	8.500	.333	0.256
15	5.4	3	.500	.600	1.000	.500	10.300	.925	0.336
16	5.6	4	.500	.600	1.000	.625	11.425	.758	0.344
17	5.7	2	.500	.800	1.000	.625	12.825	.783	0.379
18	5.8	1	.667	.800	1.000	.625	13.993	.493	0.394
19	6.0	4	.667	.800	1.000	.750	15.367	.325	0.526
20	6.4	4	.667	.800	1.000	.875	17.013	.346	0.588
21	6.5	4	.667	.800	1.000	1.000	18.867	.501	0.705
22	6.9	1	.833	.800	1.000	1.000	20.365	.199	1.119
23	7.0	2	.833	1.000	1.000	1.000	22.165	.123	1.351
24	7.6	1	1.000	1.000	1.000	1.000	24.000	0.000	1.553
							$\sum_r S_+(r) = 217$		

From the table one computes readily

$$W^2 = \frac{217.170}{24} - \frac{(25)(49)}{6(24)} = .542;$$

$D = \max\ (.158,\ .200, \cdots,\ .123,\ 0.000) = 1.583$; and

$\hat{M} = 1.$

Consequently, the decisions are

$W^2(6, 5, 5, 8)$ detector: NO

$D(6, 5, 5, 8)$ detector: NO

$\hat{M}$ detector: NO

In order to compute $S(U_o)$, $S'(\Phi)$ and $Q(n_1, \cdots, n_4; 3)$ it is convenient to construct tables 18 – 20.

Table 18. $S(U_o)$

$i =$	1	2	3	4	
	8	9	4	16	
	2	5	15	14	
	6	10	1	20	
	24	17	7	19	
	22	23	11	13	
	18			21	
				12	
				3	
R_{i+}	80	64	38	118	300

Entries are the ranks $R(X_{ij})$.

Table 19. $\hat{S}(\Phi)$

$i =$	1	2	3	4	
	-0.177	-0.148	-0.664	0.349	
	-1.354	-0.605	0.336	0.256	
	-0.517	-0.132	-2.510	0.588	
	1.553	0.379	-0.429	0.526	
	1.119	1.351	-0.084	0.220	
	0.394			0.705	
				0.014	
				-1.167	
W_{i+}	1.018	.845	-3.351	1.486	.002

Entries are the $W[R(X_{ij})]$ replacing the X_{ij}.

Table 20. $Q(n_1, \cdots, n_4; 3)$

$i =$	1	2	3	4	TOTALS
1st Block	3	1	3	1	$8 = \frac{N}{3}$
2nd Block	0	2	2	4	$8 = \frac{N}{3}$
3rd Block	3	2	0	3	$8 = \frac{N}{3}$
$\Sigma\, n_{ij}$	6	5	5	8	24

Entries are the n_{ij}.

From these last three tables one computes, respectively,

$$S(U_o) = \frac{12}{(24)(25)} \left[\frac{(80)^2}{6} + \frac{(64)^2}{5} + \frac{(38)^2}{5} + \frac{(118)^2}{8} \right] - 3(25) = 3.432$$

$$\hat{S}(\Phi) = \left[\frac{(1.018)^2}{6} + \frac{(.845)^2}{5} + \frac{(-3.351)^2}{5} + \frac{(1.486)^2}{8} \right] - \frac{(.002)^2}{24} = 2.838;$$

and $Q(6, 5, 5, 8; 3) =$

$$3 \left[\frac{3^2 + 0^2 \div 3^2}{6} + \frac{1^2 + 2^2 + 2^2}{5} + \frac{3^2 + 2^2 + 0^2}{5} + \frac{1^2 + 4^2 + 3^2}{8} \right] - 24 = 7.95$$

These detectors then decide as follows:

<u>$S(U_o)$ detector:</u> NO

<u>$S(\Phi)$ detector:</u> NO; and

<u>$Q(6, 5, 5, 8; 3)$ detector:</u> NO.

E. GOODNESS CRITERIA FOR MODEL III DETECTORS

The goodness criteria treated for Model I and Model II detectors are primarily related to one-sided alternatives as illustrated previously (see page 41.) Unfortunately, when there are three or more samples to be compared, the term "one-sided alternatives" does not have a clear meaning, and the number of reasonable interpretations is excessively large to tabulate. Consequently, no attempt will be made to apply the max-min PFD criterion or that for PFD with Rayleigh alternatives.

No statistical work comparable to that of Bell and Doksum[12] is currently available for adaptation to Model III detectors. Hence, the ALM PFD criterion will not be employed here.

However, the statistical work of Puri[18] and of Bell and Doksum[11] can be adapted to give ARE results for some of the rank-sum Model III detectors. These results are replicas of tables 6 and 14 with the appropriate changes of statistics. From tables 21 and 22 it is evident that the following adaptation of the results of Puri[18] and of Bell and Doksum[11] is valid.

Theorem 12. Under the regularity conditions of Puri[18] and for translation alternatives, one has:

(i) $A[S(H), S'(H)] = A[S(H), \hat{S}(H)] = A[S'(H), \hat{S}(H)] = 1$ for Model III detectors; and, in comparing corresponding Model I, II, and III rank-sum detectors,

(ii) $A[K(H), K(G)] = A[V(H), V(G)] = A[V'(H), V'(G)] = A[T(H), T(G)] = A[S(H), S(G)] = A[S'(H), S'(G)] = A[\hat{S}(H), \hat{S}(G)]$, for all (sufficiently regular) H and G.

Table 21. ARE's for Model III Detectors

ARE's / TRANSLATION ALTERNATIVES	$A[S(\Phi),S(U_o)] = A[S'(\Phi),S(U_o)]$ ...... $= A[\hat{S}(\Phi),\hat{S}(U_o)]$	$A[S(\Phi),S(\widetilde{H}_e)] = A[S'(\Phi),S(\widetilde{H}_e)]$ ...... $= A[\hat{S}(\Phi),\hat{S}(\widetilde{H}_e)]$	$A[S(\Phi),S(H_e)] = A[S'(\Phi),S(H_e)]$ ...... $= A[\hat{S}(\Phi),\hat{S}(H_e)]$	$A[S(\widetilde{H}_e),S(U_o)] = A[S'(\widetilde{H}_e),S(U_o)]$ ...... $= A[\hat{S}(\widetilde{H}_e),\hat{S}(U_o)]$	$A[S(H_e),S(U_o)] = A[S'(H_e),S(U_o)]$ ...... $= A[\hat{S}(H_e),\hat{S}(U_o)]$	$A[S(H_e),S(\widetilde{H}_e)] = A[S'(H_e),S(\widetilde{H}_e)]$ ...... $= A[\hat{S}(H_e),\hat{S}(\widetilde{H}_e)]$
Normal $F_o(x) = \Phi(x)$	1.05					
Uniform $F_o(x) = U_o(x)$	∞			∞	∞	1.00
Exponential $F_o(x) = H_e(x)$	∞	∞		.333		
Negative exponential $F_o(x) = \widetilde{H}_e(x) = e^x,\ x \le 0$	∞		∞		.333	
Logistic $F_o(x) = \dfrac{1}{1+e^{-x}}$	.955	1.27	1.27	.75	.75	1.00
Double exponential $F_o(x) = \begin{cases} e^x/2, & x \le 0 \\ 1-e^x/2, & x > 0 \end{cases}$	.85					
Cauchy $F_o(x) = \dfrac{1}{2} + \dfrac{1}{\pi}\tan^{-1}x$	.71					

Table 22. Statistical Distributions for Model III Detectors

Statistic	Distribution	Source of Tables, Formulas, etc. (Numbers refer to Bibliography)
$D(n_1, \cdots, n_k, 2, 1)$	(k-sample Kolmogorov-Smirnov)	34
$W^2(n_1, \cdots, n_k, 2, 1)$	(k-sample Cramér-von Mises)	34
$Q(n_1, \cdots, n_k, b)$	(Extended median) Asymptotically chi-square	28, p. 293; 21, p. 49 ff
$\hat{M}$	(Mosteller-Tukey)	33
$S(U_0)$, $S'(U_0)$	Asymptotic chi-square (Kruskal-Wallis)	21, p. 49 ff; 19, p. 290 ff
$S(\Phi)$	Asymptotic chi-square (Van de Waerden)	21, p. 49 ff; 32
$S'(\Phi)$	(Fisher-Yates) Asymptotic chi-square	21, p. 49 ff; 31, p. 66 ff
$\hat{S}(\Phi)$	Chi-square	21, p. 49 ff; 11

REVERSE SIDE BLANK

BIBLIOGRAPHY

1. Columbia University. Electrical Engineering Department Technical Report T-1/N, Nonparametric Methods For the Detection of Signals in Noise, by J. Capon, 12 March 1959

2. Birnbaum, Z. W. and Rubin, H., "On Distribution-Free Statistics," Annals of Mathematical Statistics, v.25, p.593-598, 1954

3. Bell, C. B., "On the Structure of Distribution-Free Statistics," Annals of Mathematical Statistics, v.31, p.703-709, September 1960

4. Bell, C. B., "Some Basic Theorems of Distribution-Free Statistics," Annals of Mathematical Statistics, v.35, p.150-156, March 1964

5. Bell, C. B., "A Characterization of Multisample Distribution-Free Statistics," Annals of Mathematical Statistics, v.35, p.735-738, June 1964

6. Scheffé, H., "On a Measure Problem Arising in the Theory of Non-Parametric Tests," Annals of Mathematical Statistics, v.14, p.227-233, 1943

7. Pitman, E. J. G., "Tests of Hypotheses Concerning Location and Scale Parameters," Biometrika, v.31, p.200-215, July 1939

8. Kendall, M. G. and Stuart, A., The Advanced Theory of Statistics, V.II: Inference and Relationship, Hafner, 1961

9. Chapman, D. G., "A Comparative Study of Several One-Sided Goodness-of-Fit Tests," Annals of Mathematical Statistics, v.29, p.655-674, September 1958

10. Bell, C. B. and others, "Goodness Criteria For Two-Sampled Distribution-Free Statistics," Annals of Mathematical Statistics, 1964 (Accepted for Publication)

11. Bell, C. B. and Doksum, K. A., "Some New Distribution-Free Statistics," Annals of Mathematical Statistics, 1964 (Accepted for Publication)

12. Bell, C. B. and Doksum, K. A., "Optimal One-Sample Distribution-Free Tests and Their Two-Sample Extensions," 1964 (Submitted for Publication)

13. Capon, J., "Asymptotic Efficiency of Certain Locally Most Powerful Rank Tests," Annals of Mathematical Statistics, v.32, p.88-100, March 1961

14. Chernoff, H. and Savage, I. R., "Asymptotic Normality and Efficiency of Certain Nonparametric Test Statistics," Annals of Mathematical Statistics, v.29, p.972-994, December 1958

15. Hodges, J. L. and Lehmann, E. L., "The Efficiency of Some Non-Parametric Competitors of the t-Test," Annals of Mathematical Statistics, v.27, p.324-335, 1956

16. Hodges, J. L. and Lehmann, E. L., "Comparison of the Normal Scores and Wilcoxon Tests," p.307-317 in Berkeley Symposium on Mathematical Statistics and Probability, University of California, Proceedings. Fourth, v.1: Contributions to the Theory of Statistics, University of California Press, 1961

17. Hoeffding, W., "'Optimum' Nonparametric Tests," p.83-92 in Berkeley Symposium on Mathematical Statistics and Probability, University of California, Proceedings. Second, University of California Press, 1951

18. Puri, M. L., "Asymptotic Efficiency of a Class of c-Sample Tests," Annals of Mathematical Statistics, v.35, p.102-121, March 1964

19. Birnbaum, Z. W., "Numerical Tabulation of the Distribution of Kolmogorov's Statistic for Finite Sample Size," American Statistical Association. Journal, v.47, p.425-441, September 1952

20. Marshall, A. W., "The Small Sample Distribution of nw^2_n," Annals of Mathematical Statistics, v.29, p.307-309, March 1958

21. Owen, D. B., Handbook of Statistical Tables, Addison-Wesley, 1962

22. Birnbaum, Z. W. and Tingey, F. H., "One-Sided Confidence Contours for Probability Distribution Functions," Annals of Mathematical Statistics, v.22. p.592-596, 1951

23. Darling, D. A. "The Kolmogorov-Smirnov, Cramér-Von Mises Tests," Annals of Mathematical Statistics, v.28, p.823-838, December 1957

24. Lewis, P. A., "Distribution of the Anderson-Darling Statistic," Annals of Mathematical Statistics, v.32, p.1118-1124, December 1961

25. Matusita, K., "Decision Rules, Based on the Distance, For Problems of Fit, Two Samples, and Estimation," Annals of Mathematical Statistics, v.26, p.631-640, 1955

26. Sherman, B., "A Random Variable Related to the Spacing of Sample Values," Annals of Mathematical Statistics, v.21, p.339-361, 1950

27. Kendall, M. G. and Stuart, A., The Advanced Theory of Statistics, V.I: Distribution Theory, Hafner, 1958

28. Dixon, W. J. and Massey, Jr., F. J., Introduction to Statistical Analysis, 2d ed., McGraw-Hill, 1957

29. Wilks, S. S., Mathematical Statistics, Wiley, 1962

30. Moses, L. E., "Non-Parametric Statistics For Psychological Research," Psychological Bulletin, v.49, p.122-143, 1952

31. Fisher, R. A. and Yates, F., Statistical Tables For Biological, Agricultural, and Medical Research, 3d ed., Hafner, 1948

32. van der Waerden, B. L. and Nievergelt, E., Tables for Comparing Two Samples by X-Test and Sign Test, Springer-Verlag, Berlin, 1956

33. Mosteller, F. and Tukey, J. W., "Significance Levels for a k-Sample Slippage Test," Annals of Mathematical Statistics, v.21, p.120-123, 1950

34. Kiefer, J., "K-Sample Analogues of the Kolmogorov-Smirnov and Cramér-V. Mises Tests," Annals of Mathematical Statistics, v.30, p.420-447, June 1959

Anderson, T. W. and Darling, D. A., "Asymptotic Theory of Certain 'Goodness-of-Fit' Criteria Based on Stochastic Processes," Annals of Mathematical Statistics, v.23, p.193-212, 1952

Anderson, T. W. and Darling, D. A., "A Test of Goodness of Fit," American Statistical Association. Journal, v.49, p.765-769, December 1954

Bell, C. B., Unpublished Nonparametric Class Notes, 1964

Capon, J., "A Nonparametric Technique For the Detection of a Constant Signal in Additive Noise," Institute of Radio Engineers. Wescon Convention Record, v.4, p.92-103, 18-21 August 1959

Capon, J., "Optimum Coincidence Procedures For Detecting Weak Signals in Noise," Institute of Radio Engineers. International Convention Record, v.4, p.154-156, 21-24 March 1960

Navy Electronics Lab., San Diego, Calif.
Report 1245

AUTOMATIC DISTRIBUTION-FREE STATISTICAL SIGNAL DETECTION, by C. B. Bell. 144 p., 16 Oct 64.

UNCLASSIFIED

Distribution-free statistical techniques are reviewed and evaluated for possible application to the signal detection problem. These techniques are classified into three types: (1) The noise distribution is known but the signal-plus-noise is not, (2) neither distribution is known but a pure-noise sample is available, and (3) the distributions and a pure-noise sample are all unavailable. Several of the distribution-free detectors are almost as efficient as their parametric counterparts and several are uniformly better than the parametric detectors for certain classes of Gaussian noise.

1. Signals - Detection
2. Statistical analysis

I. Bell, C. B.

SF 001 02 05, Task 6072
(NEL D1-17-5)

This card is UNCLASSIFIED

Navy Electronics Lab., San Diego, Calif.
Report 1245

AUTOMATIC DISTRIBUTION-FREE STATISTICAL SIGNAL DETECTION, by C. B. Bell. 144 p., 16 Oct 64.

UNCLASSIFIED

Distribution-free statistical techniques are reviewed and evaluated for possible application to the signal detection problem. These techniques are classified into three types: (1) The noise distribution is known but the signal-plus-noise is not, (2) neither distribution is known but a pure-noise sample is available, and (3) the distributions and a pure-noise sample are all unavailable. Several of the distribution-free detectors are almost as efficient as their parametric counterparts and several are uniformly better than the parametric detectors for certain classes of Gaussian noise.

1. Signals - Detection
2. Statistical analysis

I. Bell, C. B.

SF 001 02 05, Task 6072
(NEL D1-17-5)

This card is UNCLASSIFIED

Navy Electronics Lab., San Diego, Calif.
Report 1245

AUTOMATIC DISTRIBUTION-FREE STATISTICAL SIGNAL DETECTION, by C. B. Bell. 144 p., 16 Oct 64.

UNCLASSIFIED

Distribution-free statistical techniques are reviewed and evaluated for possible application to the signal detection problem. These techniques are classified into three types: (1) The noise distribution is known but the signal-plus-noise is not, (2) neither distribution is known but a pure-noise sample is available, and (3) the distributions and a pure-noise sample are all unavailable. Several of the distribution-free detectors are almost as efficient as their parametric counterparts and several are uniformly better than the parametric detectors for certain classes of Gaussian noise.

1. Signals - Detection
2. Statistical analysis

I. Bell, C. B.

SF 001 02 05, Task 6072
(NEL D1-17-5)

This card is UNCLASSIFIED

Navy Electronics Lab., San Diego, Calif.
Report 1245

AUTOMATIC DISTRIBUTION-FREE STATISTICAL SIGNAL DETECTION, by C. B. Bell. 144 p., 16 Oct 64.

UNCLASSIFIED

Distribution-free statistical techniques are reviewed and evaluated for possible application to the signal detection problem. These techniques are classified into three types: (1) The noise distribution is known but the signal-plus-noise is not, (2) neither distribution is known but a pure-noise sample is available, and (3) the distributions and a pure-noise sample are all unavailable. Several of the distribution-free detectors are almost as efficient as their parametric counterparts and several are uniformly better than the parametric detectors for certain classes of Gaussian noise.

1. Signals - Detection
2. Statistical analysis

I. Bell, C. B.

SF 001 02 05, Task 6072
(NEL D1-17-5)

This card is UNCLASSIFIED

INITIAL DISTRIBUTION LIST

CHIEF, BUREAU OF SHIPS
  CODE 24DC (2) CODE 210L (2)
  CODE 320
  CODE 360 (2)
  CODE 452E
  CODE 67D (2)
  CODE 684 (2)
  CODE 685A
CHIEF, BUREAU OF NAVAL WEAPONS
  DLI-3
  DLI-31
  R-56
  RUDC-2
  RUDC-11
CHIEF OF NAVAL PERSONNEL
  PERS 11B
CHIEF OF NAVAL OPERATIONS
  OP-07T
  OP-71
  OP-94G43
  OP-03EG
CHIEF OF NAVAL RESEARCH
  CODE 455
COMMANDER IN CHIEF US PACIFIC FLEET
COMMANDER IN CHIEF US ATLANTIC FLEET
COMMANDER OPERATIONAL TEST AND
  EVALUATION FORCE
DEPUTY COMMANDER OPERATIONAL TEST -
  EVALUATION FORCE, PACIFIC
COMMANDER CRUISER-DESTROYER FORCE,
  US ATLANTIC FLEET
  US PACIFIC FLEET
COMMANDER TRAINING COMMAND
  US PACIFIC FLEET
COMMANDER AMPHIBIOUS FORCE
  US PACIFIC FLEET (2)
COMMANDER SERVICE FORCE
  US ATLANTIC FLEET
DESTROYER DEVELOPMENT GROUP PACIFIC
US NAVAL AIR DEVELOPMENT CENTER
  NADC LIBRARY
US NAVAL MISSILE CENTER
  TECH. LIBRARY, CODE NO 3022
  CODE N3232
US NAVAL AIR TEST CENTER
  NANEP
US NAVAL ORDNANCE LABORATORY
  LIBRARY
US NAVAL ORDNANCE TEST STATION
  PASADENA ANNEX LIBRARY
  CHINA LAKE
FLEET COMPUTER PROGRAMMING CENTER
  ATLANTIC, TECHNICAL LIBRARY
US NAVAL WEAPONS LABORATORY
  LIBRARY
USN RADIOLOGICAL DEFENSE LABORATORY
DAVID TAYLOR MODEL BASIN
  /LIBRARY/
US NAVY MINE DEFENSE LABORATORY
USN UNDERWATER SOUND LABORATORY
  LIBRARY (3)
ATLANTIC FLEET ASW TACTICAL SCHOOL
USN MARINE ENGINEERING LABORATORY
US NAVAL RESEARCH LABORATORY
  CODE 2027
  CODE 5120
  CODE 5320
  CODE 5330
  CODE 5400
US NAVAL ORDNANCE LABORATORY
  CORONA
USN UNDERWATER SOUND REFERENCE LAB.
BEACH JUMPER UNIT ONE
BEACH JUMPER UNIT TWO
OFFICE OF NAVAL RESEARCH
  PASADENA
USN PERSONNEL RESEARCH ACTIVITY
  SAN DIEGO
US NAVAL POSTGRADUATE SCHOOL
  LIBRARY (2)
NAVY REPRESENTATIVE
  MIT LINCOLN LABORATORY
ASSISTANT SECRETARY OF THE NAVY R-D
AIR DEVELOPMENT SQUADRON ONE /VX-1/
CHARLESTON NAVAL SHIPYARD
DEFENSE DOCUMENTATION CENTER (20)
DOD RESEARCH AND ENGINEERING
  TECHNICAL LIBRARY
FEDERAL AVIATION AGENCY
  SYSTEMS RESEARCH - DEVELOPMENT SERV.
NATIONAL OCEANOGRAPHIC DATA CENTER (2)
NASA
  LANGLEY RESEARCH CENTER (3)
CHIEF, US WEATHER BUREAU (2)
CHIEF, WEATHER RADAR LABORATORY
  NORMAN, OKLAHOMA
NATIONAL SECURITY AGENCY
  C 121
  C 124
NATIONAL BUREAU OF STANDARDS
  BOULDER LABORATORIES
US BUREAU OF COMMERCIAL FISHERIES
  HONOLULU-JOHN C MARR
NAFEC LIBRARY
US ARMY MATERIAL COMMAND
ASST CHIEF OF STAFF FOR INTELLIGENCE
  US ARMY
ABERDEEN PROVING GROUND, MARYLAND (2)
US ARMY TEST - EVALUATION COMMAND (2)
US ARMY ELECTRONIC PROVING GROUND
REDSTONE SCIENTIFIC INFORMATION
  CENTER
US ARMY ELECTRONICS R-D LABORATORY
  SELRA/SR
  SELRA/SL-ADT
HARRY DIAMOND LABORATORIES
  ATTN- LIBRARY
PICATINNY ARSENAL
US ARMY RESEARCH OFFICE /DURHAM/
ELECTRONIC DEVELOPMENT ACTIVITY
  WHITE SANDS MISSILE RANGE
ARMY TRANSPORATION TERMINAL
  COMMAND, PACIFIC
US ARMY AIR DEFENSE BOARD
  SUPPORT DIVISION
EDGEWOOD ARSENAL
US ARMY ENGINEER R-D LABORATORIES
  STINFO BRANCH (2)
DEPUTY CHIEF OF STAFF, US AIR FORCE
  AFOCC-C/I
  AFRST-EL/CS
AIR DEFENSE COMMAND
  ADOOA
AIR RESEARCH AND DEVELOPMENT COMMAND
AIR UNIVERSITY LIBRARY
HQ ALASKAN AIR COMMAND (2)
STRATEGIC AIR COMMAND
AIR FORCE MISSILE TEST CENTER
  /AFMTC TECH LIBRARY - MU-135/
ROME AIR DEVELOPMENT CENTER
HQ AIR WEATHER SERVICE
WRIGHT-PATTERSON AF BASE
  AERONAUTICAL SYSTEMS DIV-ASNVEG
  ASD ASAPRD-DIST
  RTD-AVWE, ANTENNA-RADOME GROUP
  SEG/SEODE/
USAF SECURITY SERVICE
  ESD/ESA
UNIVERSITY OF CALIFORNIA
  ELECTRONICS RESEARCH LAB
UNIVERSITY OF MIAMI
  THE MARINE LAB. LIBRARY (3)
COLUMBIA UNIVERSITY
  HUDSON LABORATORIES
GEORGIA INSTITUTE OF TECHNOLOGY
  CHIEF, ELECTRONICS DIVISION
NEW YORK UNIVERSITY
  DEPT OF METEOROLOGY - OCEANOGRAPHY
UNIVERSITY OF MICHIGAN
  DIRECTOR, COOLEY ELECTRONICS LAB
  OFFICE OF RESEARCH ADMIN
  RADAR LABORATORY
  THE RADIATION LABORATORY
OHIO STATE UNIVERSITY
  ANTENNA LABORATORY
UNIVERSITY OF HAWAII
  ELECTRICAL ENGINEERING DEPT
HARVARD UNIVERSITY
  GORDON MCKAY LIBRARY
UNIVERSITY OF ILLINOIS
  ANTENNA LABORATORY
STANFORD ELECTRONICS LABORATORIES
  ATTN- DOCUMENTS LIBRARY
MASSACHUSETTS INST OF TECHNOLOGY
  RESEARCH LAB OF TECHNOLOGY
  ENGINEERING LIBRARY
MIT-LINCOLN LABORATORY
  LIBRARY, A-082
  DIVISION 3
  HEAD, RADAR DIVISION
UNIVERSITY OF NEW MEXICO
  ENGINEERING EXPERIMENT STATION
THE JOHNS HOPKINS UNIVERSITY
  APPLIED PHYSICS LABORATORY